ON THE DIFFERENCE BETWEEN CONSECUTIVE PRIMES

JAMES MAYNARD

Abstract. We show that the sum of squares of differences between consecutive primes $\sum_{p_n\le x}(p_{n+1}-p_n)^2$ is bounded by $x^{5/4+\epsilon}$ for x sufficiently large and any fixed $\epsilon>0$. This reproduces an earlier result of Peck, which the author was initially unaware of.

1. Update: 16/01/2012

The same result was obtained by Peck [23] in his thesis in 1996. The methods used here are fundamentally the same. This work does not include any new results.

2. Introduction and Context

One central topic in number theory is understanding the distribution of prime numbers. When investigating the distribution of primes, it is natural to look at the gaps between them.

We let p_n denote the n^{th} prime number, and $d_n=p_{n+1}-p_n$ denote the n^{th} prime gap.

2.1. **Average Size of Prime Gaps.**

The prime number theorem was conjectured by Gauss in 1792, and proven independently by Hadamard [8] and de la Vallée Poussin [5]. It states that

$$\pi(x)\sim\frac{x}{\log x}. \tag{1}$$

This shows that

$$\mathop{\mathbb{E}}_{x\le p_n\le 2x} d_n \sim \log x, \tag{2}$$

and so the average gap between primes of size approximately x is $\log x$.

Since $\log x$ is small in comparison with x (the size of primes we are considering), it is natural to consider how much larger d_n can be than this average. The basic intuition is that prime numbers are reasonably regular, and so the difference between consecutive primes can not be unusually ‘large’.

Supported by EPSRC Doctoral Training Grant EP/P505216/1 .

2.2. **Numerical Evidence and Heuristics.** When obtaining results for prime gaps we are usually interested in large primes (primes outside of any computable range). Computable primes are not necessarily representative of all primes. (For example, Littlewood's result that $\pi(x) > \mathrm{li}(x)$ infinitely often requires large primes. The first occurrence of this is well outside computational bounds). In particular, most results bounding the size of d_n take the form $d_n \ll f(p_n)$ with no explicit size of the implied constant. This constant would likely dominate the bound in any computable region if it was effective and calculated.

That said, it can be interesting to look at the size of prime gaps in a computable region.

N	$\max_{p_n\le N} d_n$	$\max_{p_n\le N}(\log d_n)/(\log N)$
10^1	4	0.60
10^2	8	0.45
10^3	20	0.43
10^4	36	0.39
10^5	72	0.37
10^6	114	0.34
10^7	154	0.31
10^8	220	0.29
10^9	282	0.27
10^{10}	354	0.25
10^{11}	464	0.24
10^{12}	540	0.23
10^{13}	674	0.22
10^{14}	804	0.21
10^{15}	906	0.20
10^{16}	1132	0.19

Relative to the size of the primes, the gaps between the primes of this size remain very small. Based on this very limited numerical evidence, it appears that

$$\log(d_n)/\log(p_n) \to 0.$$

This is equivalent to the statement

$$d_n \ll p_n^{\epsilon} \tag{3}$$

for any $\epsilon > 0$.

Based on numerical evidence Legendre [21] conjectured in 1798 that there is always a prime between any pair of consecutive squares. Proving this requires an estimate of the strength $d_n \le 2p_n^{1/2}$. Cramér [4] and Shanks [28] have made stronger conjectures based on probabilistic models of the primes. Although more sophisticated models give slightly different expectations of the maximal asymptotic size of d_n (as pointed out by Granville [7]), there appears no reason to disbelieve a conjecture such as

$$d_n \ll (\log p_n)^{2+\epsilon}. \tag{4}$$

Unfortunately even Legendre's conjecture seems beyond the current machinery for dealing with primes (even under the assumption of strong conjectures such as the Riemann Hypothesis). A result as strong as Cramér's conjecture appears completely impossible to prove with the available techniques.

2.3. **Large Prime Gap Bounds.** Although we cannot prove Legendre's or Cramér's conjectures, we can still obtain non-trivial bounds on the size of d_n.

Bertrand's Postulate states that there is always a prime between any integer n and $2n-2$. This was conjectured in 1845 by Bertrand [2] and proven in 1850 by Chebyshev [3]. There is therefore a prime between p_n+1 and $2p_n$, and so we must have

$$d_n \leq p_n. \tag{5}$$

Further advancements were then made from analysing the distribution of zeroes of the Riemann Zeta function. Hoheisel [17] showed that

$$d_n \ll p_n^{32999/33000}. \tag{6}$$

The exponent of p_n in the right hand side has been repeatedly reduced by different authors including Heilbronn [16], Tchudakoff [29], Ingham [20] and Huxley [18]. These improvements were largely down to the development of more sophisticated methods to analyse the distribution of the zeroes of $\zeta(s)$. The most recent result is due to Baker, Harman and Pintz [1], which shows that

$$d_n \ll p_n^{21/40}. \tag{7}$$

Better results can be obtained if we assume conjectures about the Riemann Zeta function.

Riemann famously conjectured that all the non-trivial zeroes of $\zeta(s)$ have real part $1/2$. Cramér [4] showed that assuming the Riemann Hypothesis

$$d_n \ll p_n^{1/2} \log p_n. \tag{8}$$

The density hypothesis states that the number of zeroes $N(\sigma, T)$ of $\zeta(s)$ with absolute value of imaginary part less than T and real part greater than σ satisfies $N(\sigma, T) \ll T^{2(1-\sigma)} \log^A T$ for some constant A. This follows from the Riemann Hypothesis or the Lindelöf Hypothesis. Assuming this weaker hypothesis one can prove

$$d_n \ll p_n^{1/2+\epsilon} \tag{9}$$

for any $\epsilon > 0$. Both of these conditional results would therefore show that there is always a prime in the interval $[x, x + x^{1/2+\epsilon}]$ for x sufficiently large.

2.4. **Lower Bounds on Large Prime Gaps.** One can construct sequences of consecutive composite integers to explicitly demonstrate large gaps between primes. For example,

$$j + \prod_{i \leq n} p_i \tag{10}$$

is clearly composite for $2 \leq j \leq p_n$. This (and small refinements) show that $d_n \geq C \log n$ for some constant C.

Westzynthius [31] showed in 1931 that by carefully sieving certain primes one can have gaps between primes which are larger than any constant multiple of the average gap $\log p_n$. Erdős [6] and Rankin [25] subsequently improved the size of this lower bound on d_n using similar ideas. The best current result is due to Pintz [24] which states that for infinitely many integers n we have

$$d_n \geq (2e^\gamma + o(1)) \frac{(\log n)(\log\log n)(\log\log\log\log n)}{(\log\log\log n)^2}. \tag{11}$$

Note that this lower bound is only slightly larger than the average bound $\log n$, and is less than the upper bound of $\log^2 n$ predicted by Cramér's conjecture.

2.5. **Frequency of Large Prime Gaps.** The Results of Section 2.1 give a precise asymptotic value of the L^1 norm of d_n from the Prime Number Theorem.

The results in Section 2.3 give bounds on the L^∞ norm of d_n, but fall short of what the expected bound on gaps between primes should be, even with the assumption of the Riemann Hypothesis. It seems with the current technology we cannot hope to prove anything close to the true size of the L^∞ bound.

It is therefore natural to look at the L^2 norm of d_n. Even if we cannot show that unusually large gaps do not occur, we can hope to show that the vast majority of prime gaps are much smaller and that large gaps, should they exist, are infrequent.

Selberg [27] proved, assuming the Riemann Hypothesis, that

$$\sum_{p_n \le x} d_n^2 \ll x(\log x)^3. \tag{12}$$

In particular, this shows that almost all intervals $[x, x + (\log x)^{2+\epsilon}]$ contain a prime, and that the root mean square gap between primes is $\ll (\log x)^2$. These results therefore show (assuming the Riemann hypothesis) that at least a majority of gaps satisfy bounds similar to those predicted by Cramér's conjecture.

Yu [32] improved a result of Heath-Brown [11] to prove, assuming the Lindelöf Hypothesis, that

$$\sum_{p_n \le x} d_n^2 \ll x^{1+\epsilon} \tag{13}$$

for any $\epsilon > 0$. Both of these results show that almost all intervals $[x, x + x^\epsilon]$ contain a prime. Thus a claim '$d_n \ll p_n^\epsilon$' would at least hold for almost all prime gaps.

The best unconditional L^2 result thus far is due to Heath-Brown [12], who proved that

$$\sum_{p_n \le x} d_n^2 \ll x^{23/18+\epsilon}. \tag{14}$$

This shows that $\sum_{d_n \ge x^a} d_n \ll x^{23/18-a+\epsilon}$. It immediately follows that $d_n \ll p_n^{23/36+\epsilon}$ and almost all intervals $[x, x+x^{5/18+\epsilon}]$ contain a prime. Although both of these can be improved with alternative methods, we note that the exponent of $5/18+\epsilon$ is much smaller than the Riemann Hypothesis bound of $1/2+\epsilon$, and so being able to ignore a small number of possible large differences makes the problem much more tractable.

These results should be compared with the lower bound obtained by the Cauchy-Schwartz inequality and the prime number theorem, which gives

$$\sum_{p_n \le x} d_n^2 \gg x \log x. \tag{15}$$

3. Main Result

We aim to improve on Heath-Brown's result [12] and investigate the values of ν for which we can show

$$\sum_{p_n \le x} d_n^2 \ll x^{1+\nu+\epsilon} \tag{16}$$

for any $\epsilon > 0$.

We do this by obtaining L^2, L^4 and L^∞ bounds on the Chebyschev function $\psi(x) = \sum_{n \le x} \Lambda(n)$ in intervals of size τ.

In particular, we wish to prove:

Theorem 3.1.
$$\sum_{p_n \le x} d_n^2 \ll x^{5/4+\epsilon}$$
for any $\epsilon > 0$.

By dyadic subdivision and replacing ϵ by a finite multiple, we see it is sufficient to prove the following proposition.

Proposition 3.2. *For* $0 \le \tau \le x$ *we have*
$$\sum_{\substack{4x/\tau \le d_n \le 8x/\tau \\ x \le p_n \le 2x}} d_n^2 \ll x^{5/4+10\epsilon}$$
for any $\epsilon > 0$

4. Initial Argument

Proposition 3.2 holds trivially for $d_n \ll x^{1/4+\epsilon}$ or (by the result of Baker, Harman and Pintz [1][Theorem 1]) for $d_n \gg x^{21/40}$. Thus we only need to consider

$$x^{19/40} \le \tau \le x^{3/4-\epsilon}. \tag{17}$$

We follow essentially exactly the same method as Heath-Brown in [11] in this section, except that we use Perron's formula to get an estimate for $\psi(x)$ in terms of Dirichlet polynomials instead of zeroes of $\zeta(s)$. (An idea suggested by Heath-Brown in [14]). It is the greater control which we get from using this setup which enables us to improve the exponent from $23/18$ to $5/4$.

4.1. **A Combinatorial Identity and Perron's Formula.**

We start with the identity:

$$-\frac{\zeta'}{\zeta}(s) = -\frac{\zeta'}{\zeta}(s)(1-M_x(s)\zeta(s))^k + \sum_{j=1}^{k}(-1)^j\binom{k}{j}M_x(s)^j\zeta(s)^{j-1}\zeta'(s), \tag{18}$$

where

$$k \in \mathbb{Z}^+ \text{is a positive constant,} \qquad M_x(s) = \sum_{n\le(3x)^{1/k}}\mu(n)n^{-s}. \tag{19}$$

We will later (equation (143)) choose $k = 60$, since this is sufficient for our purposes.

By our choice of M_x the term $(1-M_x(s)\zeta(s))^k\zeta'(s)/\zeta(s)$ makes no contribution to the coefficient of n^{-s} for $n \le 3x$. Hence equating coefficients of n^{-s} of both sides for $n \le 3x$ gives

$$\Lambda(n) = \sum_{j=1}^{k}(-1)^j\binom{k}{j}K^{(j)}(n) = \sum_{j=1}^{k}c_jK^{(j)}(n) \tag{20}$$

where

$$K^{(j)}(n) = \sum_{\substack{\prod_1^{2j} n_i = n\\ n_i\le(3x)^{1/k} \text{ for } i\le j}}\mu(n_1)\dots\mu(n_j)\log n_{2j}. \tag{21}$$

We split $K^{(j)}(n)$ into dyadic intervals for each n_i, therefore expressing $\Lambda(n)$ as a linear combination of $O(\log^{2k}x)$ sums of the form

$$J_{N_1,N_2,\dots,N_{2k}}(n) = \sum_{\substack{n_i\in(N_i,2N_i]\forall i\\ \prod_i n_i=n}}\mu(n_1)\dots\mu(n_k)\log n_{2k}. \tag{22}$$

We note that

$$N_i \le (3x)^{1/k} \tag{23}$$

for $i \le k$. We account for the cases when $j < k$ by setting $N_i = 1/2$ (and so $n_i = 1$) for the 'extra' variables.

We now put

$$S_i(s) = \begin{cases}\sum_{N_i<n_i\le 2N_i}\mu(n_i)n_i^{-s}, & i\le k\\ \sum_{N_i<n_i\le 2N_i}n_i^{-s} & k<i<2k\\ \sum_{N_i<n_i\le 2N_i}(\log n_i)n_i^{-s} & i=2k\end{cases} \tag{24}$$

and consider the Dirichlet polynomial

$$\sum a_n n^{-s} = \sum_{j=1}^{k} c_j \sum_{\substack{(N_i)_1^{2k} \\ N_i \le (3x)^{1/k} \text{ for } i \le k \\ 2^{-2k} x \le \prod_1^{2k} N_i \le 3x \\ N_i = 1/2 \text{ if } j < i \le k \text{ or } k+j \le i < 2k}} S_1(s) S_2(s) \dots S_{2k}(s). \tag{25}$$

By the above identity, for $x \le n \le 3x$ we have

$$a_n = \Lambda(n). \tag{26}$$

In particular, for $x \le y \le 2x$ and $\tau \ge 2$

$$\psi(y + y/\tau) - \psi(y) = \sum_{y < n \le y + y/\tau} a_n. \tag{27}$$

We separate out the case when one of the $N_i > x^{19/20}$, since such very long polynomials require a slightly different treatment.

Thus

$$\sum a_n n^{-s} = \sum f_n n^{-s} + \sum g_n n^{-s} \tag{28}$$

where

$$\sum f_n n^{-s} = \sum_{j=1}^{k} c_j \sum_{\substack{(N_i)_1^{2k} \\ N_i \le (3x)^{1/k} \text{ for } i \le k \\ 2^{-2k} x \le \prod_1^{2k} N_i \le 3x \\ N_i = 1/2 \text{ if } j < i \le k \text{ or } j+k \le i < 2k \\ N_i > x^{19/20} \text{ for some } i}} S_1(s) S_2(s) \dots S_{2k}(s), \tag{29}$$

and

$$\sum g_n n^{-s} = \sum_{j=1}^{k} c_j \sum_{\substack{(N_i)_1^{2k} \\ N_i \le (3x)^{1/k} \text{ for } i \le k \\ 2^{-2k} x \le \prod_1^{2k} N_i \le 3x \\ N_i = 1/2 \text{ if } j < i \le k \text{ or } j+k \le i < 2k \\ N_i \le x^{19/20} \text{ for all } i}} S_1(s) S_2(s) \dots S_{2k}(s). \tag{30}$$

We first consider $\sum f_n$. We separate the exceptionally long polynomial $S_{i_0}(s)$, and just consider the remaining product of polynomials as a single polynomial. Since $N_i \le (3x)^{1/k} \le x^{19/20}$ for $i \le k$ we must have $i_0 > k$ and so the exceptional polynomial must have all coefficients 1 or $\log n$. We will assume that $i_0 \neq 2k$, so all the coefficients are identically 1. The alternative case $i_0 = 2k$ may be handled similarly.

$$\sum_{y < n \le y + y/\tau} f_n = \sum_{j=1}^{k} \sum_{N_{i_0} > x^{19/20}} \sum_{2^{-2k} x / N_{i_0} \le M \le 3x / N_{i_0}} \sum_{\substack{m,n \\ y < mn \le y + y/\tau \\ M < m \le 2^{2k-1} M \\ N_{i_0} < n \le 2N_{i_0}}} b_m^{(j)} \tag{31}$$

for some coefficients $b_m^{(j)} \ll x^{\epsilon}$.

We let

$$\mathcal{B}_\tau = \left\{ z : x \le z \le 2x, mN_{i_0} > z > \frac{mN_{i_0}}{1 + 1/\tau} \text{ for some } m, N_{i_0} \right\} \tag{32}$$

and consider separately $y\notin\mathcal{B}_\tau$ and $y\in\mathcal{B}_\tau$. We note that

$$\text{meas}(\mathcal{B}_\tau)\ll\sum_{m,N_{i_0}}\frac{mN_{i_0}}{\tau}\ll\sum_{m,N_{i_0}}\frac{x}{\tau}\ll\frac{x^{21/20+\epsilon}}{\tau}. \tag{33}$$

Thus in particular $[x,2x]-\mathcal{B}_\tau\neq\emptyset$ and $\mathcal{B}_\tau$ represents only a small subset of y with $x\leq y\leq 2x$.

If $y\notin\mathcal{B}_\tau$ then for any m we have

$$\#\{n:\frac{y}{m}<n\leq\frac{y(1+1/\tau)}{m},N_{i_0}<n\leq 2N_{i_0}\}=\begin{cases}\frac{y}{\tau m}+O(1), & mN_{i_0}<y\leq 2mN_{i_0}\\ 0, & \text{otherwise}\end{cases} \tag{34}$$

Thus the sum over n is over $y/(\tau m)+O(1)$ terms if $mN_{i_0}<y<2mN_{i_0}$ or is empty. Hence

$$\begin{aligned}\sum_{y<n\leq y+y/\tau}f_n&=\mathbf{1}_{\mathcal{B}_\tau^C}(y)\sum_{\substack{m,j,M,N_{i_0}\\ mN_{i_0}<y<2mN_{i_0}}}\left(\frac{b_m^{(j)}y}{\tau m}+O(x^\epsilon)\right)+\mathbf{1}_{\mathcal{B}_\tau}(y)\sum_{y<n\leq y+y/\tau}f_n\\ &=\mathbf{1}_{\mathcal{B}_\tau^C}(y)\frac{A_1(y)}{\tau}+O\left(\sum_M Mx^\epsilon\right)+\mathbf{1}_{\mathcal{B}_\tau}(y)\left(\sum_{y\leq n\leq y+y/\tau}f_n\right),\end{aligned} \tag{35}$$

where we have defined

$$A_1(y)=\sum_{\substack{m,j,M,N_{i_0}\\ mN_{i_0}<y<2mN_{i_0}}}\frac{b_m^{(j)}y}{m}. \tag{36}$$

We note that $A_1(y)$ is independent of τ. Since $M\ll x/N_{i_0}\ll x^{1/20}$ we have

$$\begin{aligned}\sum_{y<n\leq y+y/\tau}f_n&=\frac{A_1(y)}{\tau}+O(x^{1/20+2\epsilon})+\mathbf{1}_{\mathcal{B}_\tau}(y)O\left(\sum_{y\leq n\leq y+y/\tau}f_n+A_1(y)/\tau\right)\\ &=\frac{A_1(y)}{\tau}+E_1+E_2.\end{aligned} \tag{37}$$

where $E_1=O(x^{1/19})$ and $E_2=0$ when $y\notin\mathcal{B}_\tau(y)$.

This gives us a 'main term' $A_1(y)/\tau$, which we will estimate in Lemma 4.1, and two error terms E_1 and E_2. E_1 is always small, and so causes no problems. E_2 can only be large when $y\in\mathcal{B}_\tau$, which is a suitably small set to cause us no problems.

We now consider $\sum g_n n^{-s}$. To ease notation we put

$$S(s)=\prod_{i=1}^{2k}S_i(s), \tag{38}$$

and we let the unlabelled sum $\sum$ represent the sum

$$\sum_{j=1}^{k}c_j\sum_{\substack{(N_i)_1^{2k}\\ N_i\leq(3x)^{1/k}\text{ for }i\leq k\\ 2^{-2k}x\leq\prod_1^{2k}N_i\leq 3x\\ N_i=1/2\text{ if }j<i\leq k\text{ or }j+k\leq i<2k\\ N_i\leq x^{19/20}\text{ for all }i}} \tag{39}$$

which appears in the right hand side of (30).

Perron's formula states that for $T > 2$, $x > 0$, $x \neq 1$ and $1 < \sigma \leq 2$ we have

$$\frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT}\frac{x^s}{s}ds = H(x) + O\left(\frac{x^\sigma}{T|\log x|}\right), \tag{40}$$

where $H(x) = 0$ for $x < 1$ and $H(x) = 1$ for $x > 1$.

Using Perron's formula and putting $c = 1 + 1/\log y$:

$$\begin{aligned}\sum_{y<n\leq y+y/\tau} g_n &= \frac{1}{2\pi i}\int_{c-iT_0}^{c+iT_0}\frac{y^s}{s}\left(\left(1+\frac{1}{\tau}\right)^s - 1\right)\left(\sum S(s)\right)ds + E_3 \\ &= \frac{1}{2\pi i}\int_{c-iT_1}^{c+iT_1}\frac{y^s}{s}\left(\left(1+\frac{1}{\tau}\right)^s - 1\right)\left(\sum S(s)\right)ds + E_3 + E_4 \\ &= \frac{1}{2\pi i}\int_{c-iT_1}^{c+iT_1}\frac{y^s}{\tau}\left(\sum S(s)\right)ds + E_3 + E_4 + E_5 \\ &= \frac{A_2(y)}{\tau} + E_3 + E_4 + E_5. \end{aligned} \tag{41}$$

Here

$$A_2(y) = \frac{1}{2\pi i}\int_{c-iT_1}^{c+iT_1} y^s\left(\sum S(s)\right)ds, \tag{42}$$

$$E_3 = E_3(y,\tau) = O\left(\frac{y\log^2 y}{T_0} + \log y\right), \tag{43}$$

$$E_4 = E_4(y,\tau) = O\left(\left|\int_{c+iT_1}^{c+iT_0} y^s C_1(s)\left(\sum S(s)\right)ds\right|\right), \tag{44}$$

$$E_5 = E_5(y,\tau) = O\left(\left|\int_{c-iT_1}^{c+iT_1} y^s C_2(s)\left(\sum S(s)\right)ds\right|\right), \tag{45}$$

$$C_1(s) = \frac{1}{s}\left(\left(1+\frac{1}{\tau}\right)^s - 1\right), \tag{46}$$

$$C_2(s) = \frac{1}{s}\left(\left(1+\frac{1}{\tau}\right)^s - 1 - \frac{s}{\tau}\right). \tag{47}$$

We note that

$$A_2(y) \text{ is independent of } \tau, \tag{48}$$

$$C_1(s) \ll \frac{1}{\tau}, \tag{49}$$

$$C_2(s) \ll \frac{|s|}{\tau^2}. \tag{50}$$

Therefore we have a 'main term' $A_2(y)/\tau$ and error terms E_3, E_4 and E_5. We will show that E_3 and E_5 are small, and so do not cause any problems in Lemma 4.1 below. If we can show that E_4 is only large on a small set, then we will have a suitably accurate estimate of $\psi(y + y/\tau) - \psi(y)$.

Putting together (37) and (41), using (27) and (28), and setting $A(y) = A_1(y) + A_2(y)$ we get

$$\psi(y + y/\tau) - \psi(y) = \frac{A(y)}{\tau} + E_1 + E_2 + E_3 + E_4 + E_5. \tag{51}$$

Lemma 4.1. *For $T_0 = \tau(\log y)^3$, $T_1 = y^{1/8}$ and $x^{1/3} \leq \tau \leq x^{3/4}$ we have*

$$(i) \qquad E_1, E_3, E_5 = o\left(\frac{y}{\tau}\right)(ii) \qquad A(y) \sim y \text{ for } y \notin \mathcal{B}_{y^{1/3}}.$$

Proof. (i): Estimate of E_1, E_3, E_5.

Since $\tau \leq x^{3/4}$ and $x \leq y$, we have

$$E_1 \ll x^{1/19} = o\left(\frac{y}{\tau}\right). \tag{52}$$

Since $\tau \leq y^{3/4}$ and $T_0 = \tau(\log y)^3$, we have

$$E_3 = O\left(\frac{y(\log y)^2}{T_0}\right) = O\left(\frac{y(\log y)^2}{\tau(\log y)^3} + \log y\right) = o\left(\frac{y}{\tau}\right). \tag{53}$$

We have

$$|S_i(c+it)| \leq \sum_{N_i < n_i \leq 2N_i} (\log n_i) n_i^{-c} \ll \log y \tag{54}$$

for all i. Since S is a product of the S_i we have

$$|S(c+it)| \ll y^{\epsilon}. \tag{55}$$

Thus, since $T_1 = y^{1/8}$ and $\tau \geq y^{1/3}$ and $C_2(s) \ll |s|\tau^{-2}$ (by (50)) and $\sum$ is a sum over $\ll y^{\epsilon}$ terms, we have that

$$\begin{aligned} E_5 &\ll \int_{c-iT_1}^{c+iT_1} y^c\, |C_2(s)| \left|\sum S(s)\right| |ds| \\ &\ll \frac{y^{c+2\epsilon} T_1^2}{\tau^2} \\ &\ll \frac{y^{11/12+2\epsilon}}{\tau} \\ &= o\left(\frac{y}{\tau}\right). \end{aligned} \tag{56}$$

(ii): Estimate of $A(y)$.

The idea for estimating $A(y)$ is as follows. For $\tau = \tau_0$ with τ_0 'small' we have that $\psi(y + y/\tau) - \psi(y) \sim y/\tau$. We have that E_1, E_3, E_5 are all small relative to this, and that E_2 is zero outside $\mathcal{B}_\tau$. Therefore, provided we can show E_4 is small for this value of τ, we can bound $A(y)$ from below when y is not in $\mathcal{B}_{\tau_0}$ (which covers almost all values of y). Since $A(y)$ is independent of τ, this bound holds for any size of τ, giving the result. We proceed to make this precise.

Huxley's Theorem [18] states that

$$\psi(a+b) - \psi(a) \sim b, \tag{57}$$

for $b > a^{7/12+\epsilon}$.

Using Huxley's Theorem taking $a = y$, $b = y/\tau$ and $\tau = y^{1/3}$ we obtain

$$\psi(y + y/\tau) - \psi(y) \sim y\tau^{-1}. \tag{58}$$

For this value of τ we still have

$$E_1, E_3, E_5 = o(y\tau^{-1}). \tag{59}$$

By Heath-Brown [14][Lemma 3] we have

$$\int_T^{2T} |S_1(1/2+it)\dots S_{2k}(1/2+it)|dt \ll x^{1/2}(\log x)^{-12} \tag{60}$$

uniformly for $\exp((\log x)^{1/3}) \leq T \leq x^{5/12-\epsilon}$.

A precisely analogous argument yields

$$\int_T^{2T} |S_1(c+it)\dots S_{2k}(c+it)|dt \ll x^{1-c}(\log x)^{-A} \tag{61}$$

for any constant $A > 0$ and uniformly for $\exp((\log x)^{1/3}) \leq T \leq x^{5/12-\epsilon}$. We choose $A = 2k+2$ $(= 122)$ since this will be sufficient for our purposes.

When $\tau = y^{1/3}$ we have $\exp((\log x)^{1/3}) \leq y^{1/8} = T_1$ and $T_0 = \tau(\log y)^3 \leq x^{5/12-\epsilon}$. We can therefore use (61) uniformly for $T \in [T_1, T_0]$.

Thus, since $|C_1(s)| \ll \tau^{-1}$ (by (49)), we have

$$\begin{aligned}
E_4 &\ll \left| \int_{c+iT_1}^{c+iT_0} y^s C_1(s) \sum S(s) ds \right| \\
&\ll \frac{y^c}{\tau} \sum \int_{c+iT_1}^{c+iT_0} |S(s)|\, ds \\
&\ll \frac{y^c(\log y)}{\tau} \sum \sup_{T\in[T_0,T_1]} \int_T^{2T} |S_1(c+it)\dots S_{2k}(c+it)|\, dt \\
&\ll \frac{y}{(\log y)^{2k+1}\tau} \sum 1. \qquad (62)
\end{aligned}$$

Since the sum is over $O((\log y)^{2k})$ terms, this gives

$$E_4 = o\left(\frac{y}{\tau}\right). \tag{63}$$

Thus for $\tau = y^{1/3}$ we have

$$E_1, E_3, E_4, E_5 = o(y\tau^{-1}), \qquad \psi(y + y\tau^{-1}) - \psi(y) \sim y\tau^{-1}. \tag{64}$$

Moreover, $E_2 = 0$ for $y \notin \mathcal{B}_{y^{1/3}}$ when $\tau = y^{1/3}$.

Hence for $\tau = y^{1/3}$ and $y \notin \mathcal{B}_{y^{1/3}}$

$$A(y) \sim y. \tag{65}$$

Since $A(y)$ is independent of τ, this must hold for all values of τ. □

Thus

$$\psi(y + y/\tau) - \psi(y) - \frac{y}{\tau} = E_2 + E_4 + \mathbf{1}_{\mathcal{B}_{y^{1/3}}}(y) O\left(\frac{A(y)}{\tau}\right) + o\left(\frac{y}{\tau}\right). \tag{66}$$

We let $E_6 = E_2 + \mathbf{1}_{\mathcal{B}_{y^{1/3}}}(y)A(y)/\tau$. Since $\mathcal{B}_{y^{1/3}} \supset \mathcal{B}_\tau$ for $\tau \geq y^{1/3}$ we see that $E_6 = 0$ if $y \notin \mathcal{B}_{y^{1/3}}$. Therefore

$$\psi(y + y/\tau) - \psi(y) - \frac{y}{\tau} = E_4 + E_6 + o\left(\frac{y}{\tau}\right), \tag{67}$$

where $E_6 = 0$ if $y \notin \mathcal{B}_{y^{1/3}}$.

By definition of ψ, we also have

$$\psi(y+y/\tau)-\psi(y)=\sum_{\substack{k,p \text{ prime}\\ y\le p^k\le y+y/\tau}}\log p. \tag{68}$$

The key point is that if there are no primes in the interval $[y,y+y/\tau]$ then there are no terms with $k=1$. Hence

$$\begin{aligned}\psi(y+y/\tau)-\psi(y)&\le\sum_{2\le k\le\log y}\ \sum_{y^{1/k}\le p\le y^{1/k}+(y/\tau)^{1/k}}\log y\\ &\ll(\log y)^2(y/\tau)^{1/2}\\ &=o\left(\frac{y}{\tau}\right).\end{aligned}\tag{69}$$

Thus if there are no primes in the interval $[y,y+y/\tau]$ then the left hand side of (67) is $\gg y\tau^{-1}$. This means that $E_4+E_6\gg y\tau^{-1}$. The term E_6 is only non-zero on $\mathcal{B}_{y^{1/3}}$, which is a small set, and so cannot be large frequently. Moreover, E_4 can only be large when $\sum S(c+it)$ is large, and we can show this does not happen too often by estimates on the frequency with which Dirichlet Polynomials can take large values. Thus we can show that the interval $[y,y+y/\tau]$ rarely contains no primes.

We split the sum $\sum S_1\dots S_{2k}$ up into subsums dependent on the size of each of the S_i, to show that E_4 cannot be large often.

We put

$$\mathcal{S}=\mathcal{S}(\sigma_1,\dots,\sigma_{2k}):=\left\{m\in\mathbb{Z}:N_i^{-c+\sigma_i}\le\sup_{t\in[m,m+1]}|S_i|\le 2N_i^{-c+\sigma_i}\forall i\right\}\tag{70}$$

for each $\sigma_i\in\left\{1,1-\frac{\log 2}{\log N_i},1-\frac{2\log 2}{\log N_i},\dots,-\frac{\log x}{\log N_i}\right\}$. We let $\mathcal{S}_0$ cover the remaining values of m, so $\mathcal{S}_0=\{m\in\mathbb{Z}:\sup_{t\in[m,m+1]}|S_i|\le N_i^{-c}x^{-1}$ for some $i\}$.

We let $\sum_{(\sigma_i)}$ represent the sum over all the $O((\log x)^{2k})$ values of $(\sigma_i)_1^{2k}$.

Hence splitting E_4 into terms corresponding to the choices of (σ_i) we get

$$\begin{aligned}E_4&\ll\left|\int_{c+iT_1}^{c+iT_0}y^sC_1(s)\left(\sum S(s)\right)ds\right|\\ &\ll\sum_{(\sigma_i)}\sum\left|\sum_{m\in\mathcal{S}\cap[T_0,T_1]}\int_{c+im}^{c+i(m+1)}y^sC_1(s)S(s)ds\right|\\ &\quad+\sum\ \sum_{m\in\mathcal{S}_0\cap[T_1,T_0]}\int_{c+im}^{c+i(m+1)}|y^sC_1(s)S(s)|\,|ds|.\end{aligned}\tag{71}$$

The sum $\sum$ is over $O((\log x)^{2k})$ terms, $|C_1(s)|\ll\tau^{-1}$ (by (49)) and for $m\in\mathcal{S}_0$ we have $|S(s)|\le x^{-1}$. Therefore the last term is

$$\ll(\log x)^{2k}T_0x\tau^{-1}x^{-1}=o\left(\frac{x}{\tau}\right).\tag{72}$$

We split the range of integration of the first term into $O(\log(1 + T_0/T_1))$ dyadic intervals. This gives

$$E_4 \ll \log(1 + T_0/T_1) \sup_{T\in[T_1,T_0]} \sum_{(\sigma_i)} \sum \left| \sum_{m\in\mathcal{S}\cap[T,2T]} \int_{c+im}^{c+i(m+1)} y^s C_1(s) S(s) ds \right|$$
$$+ \sum_{(\sigma_i)} \sum \int_{T_1/2}^{T_1} y^c \tau^{-1} |S(c + it)| dt + o\left(\frac{x}{\tau}\right). \tag{73}$$

We put

$$E((N_i), (\sigma_i); y, T) = \left| \sum_{m\in\mathcal{S}\cap[T,2T]} \int_m^{m+1} y^{c+it} C_1(c + it) S(c + it) dt \right|. \tag{74}$$

We note that the sum $\sum_{(\sigma_i)} \sum$ is a sum over $O((\log x)^{4k})$ terms. Thus the first term on the right hand side of (73) is

$$\ll (\log x)^{4k+1} \sup_{(N_i),(\sigma_i),T\in[T_1,T_0]} E((N_i), (\sigma_i); y, T) \tag{75}$$

where $(\sigma_i)_1^{2k}$ and $(N_i)_1^{2k}$ are constrained by

$$(i) : \sigma_i \le 1 \quad \forall i, \tag{76}$$

$$(ii) : x \ll \prod_{i=1}^{2k} N_i \ll x, \tag{77}$$

$$(iii) : N_i \le (3x)^{1/k} \text{ if } i \le k, \tag{78}$$

$$(iv) : N_i \le x^{19/20} \quad \forall i. \tag{79}$$

Since $T_1 = y^{1/8}$ we can use (61) with $A = 4k + 1$ to bound the second term on the right hand side of (73). This gives

$$\sum_{(\sigma_i)} \sum \int_{T_1/2}^{T_1} y^c \tau^{-1} |S(c + it)| dt \ll (\log x)^{4k} y^c \tau^{-1} y^{1-c} (\log x)^{-4k-1}$$
$$= o\left(\frac{y}{\tau}\right). \tag{80}$$

Since $\tau \le x^{3/4}$ we can bound the third term on the right hand side of (73) trivially.

$$(\log x)^{2k} T_0 x \tau^{-1} x^{-1} \ll (\log x)^{2k+3}$$
$$= o\left(\frac{y}{\tau}\right). \tag{81}$$

Putting this together, we obtain:

$$\left| \psi(y + y/\tau) - \psi(y) - \frac{y}{\tau} \right|$$
$$\ll (\log x)^{4k+1} \sup_{(N_i),(\sigma_i),T\in[T_1,T_0]} E((N_i), (\sigma_i); y, T) + E_6 + o\left(\frac{x}{\tau}\right). \tag{82}$$

We now want to show that there cannot be many large gaps between primes by showing that the L^2, L^4 and L^∞ norms of $E((N_i), (\sigma_i); y, T)$ cannot all be simultaneously large.

4.2. **The Basic Lemma.**

Lemma 4.2. *We have*

$$|\psi(y+y/\tau)-\psi(y)-y/\tau| \ll (\log x)^{4k+1} \sup_{(N_i),(\sigma_i),T\in[T_1,T_0]} E((N_i),(\sigma_i);y,T)+E_6+o\left(\frac{y}{\tau}\right) \tag{83}$$

where the supremum is constrained by (76),(77),(78),(79) *and* $E((N_i),(\sigma_i);y,T)$ *satisfies, for any* $\epsilon>0$*:*

$$E((N_i),(\sigma_i);y,T) \ll x_1^{\sigma}\tau^{-1}R(T), \tag{84}$$

$$\int_x^{2x} |E((N_i),(\sigma_i);y,T)|^2\,dy \ll x_1^{1+2\sigma+\epsilon}\tau^{-2}R(T), \tag{85}$$

$$\int_x^{2x} |E((N_i),(\sigma_i);y,T)|^4\,dy \ll x_1^{1+4\sigma+\epsilon}\tau^{-4}R^*(T). \tag{86}$$

Here

$$\mathcal{S}=\mathcal{S}(\sigma_1,\dots,\sigma_{2k})=\left\{m : N^{-c+\sigma_i}\le \sup_{t\in[m,m+1]}|S_i|\le 2N^{-c+\sigma_i}\forall i\right\}, \tag{87}$$

$$\mathcal{S}^*=\mathcal{S}^*(\sigma_1,\dots,\sigma_{2k})=\left\{(m_1,m_2,m_3,m_4)\in\mathcal{S}^4 : m_1+m_2=m_3+m_4\right\}, \tag{88}$$

$$R(T)=R(T,\sigma_1,\dots,\sigma_{2k})=\#\left(\mathcal{S}\cap[T,2T]\right), \tag{89}$$

$$R^*(T)=R^*(T,\sigma_1,\dots,\sigma_{2k})=\#\left(\mathcal{S}^*\cap[T,2T]^4\right), \tag{90}$$

$$x_1=\prod_{i=1}^{2k}N_i, \tag{91}$$

$$x_1^{\sigma}=\prod_i N_i^{\sigma_i}. \tag{92}$$

Proof. The Proof follows exactly the same lines as that of Heath-Brown in [10] and [11] but using Dirichlet polynomials instead of zeroes of $\zeta(s)$.

We note that

$$x\ll x_1=\prod_{i=1}^{2k}N_i\ll x, \qquad x\ll y\ll x. \tag{93}$$

We will find it slightly more convenient to work with x_1 rather than x in our later arguments, and so we introduce it now.

We recall that $|C_1(s)|\ll\tau^{-1}$ (by (49)) and that $S(c+it)\ll\prod_{i=1}^{2k}N_i^{c-\sigma_i}=x_1^{\sigma-c}$ for $t\in[m,m+1]$ and $m\in\mathcal{S}$.

(i): L^∞ estimate.

$$\begin{aligned}
E((N_i),(\sigma_i);y,T) &= \left|\sum_{m\in\mathcal{S}\cap[T,2T]}\int_m^{m+1} y^{c+it}C_1(c+it)S(c+it)dt\right| \\
&\ll \sum_{m\in\mathcal{S}\cap[T,2T]}\int_m^{m+1} y^c\tau^{-1}x_1^{-c+\sigma}dt \\
&\ll x_1^\sigma\tau^{-1}\sum_{m\in\mathcal{S}\cap[T,2T]}1 \\
&\ll x_1^\sigma\tau^{-1}R(T). \qquad (94)
\end{aligned}$$

(ii): L^2 estimate.

$$\begin{aligned}
&\int_x^{2x}|E(N_i),(\sigma_i);y,T|^2dy \\
&= \sum_{m_1,m_2\in\mathcal{S}\cap[T,2T]}\int_{c+im_1}^{c+i(m_1+1)}\int_{c+im_2}^{c+i(m_2+1)} \\
&\qquad \left(\int_x^{2x}y^{s_1+\overline{s}_2}dy\right)C_1(s_1)S(s_1)\overline{C_1(s_2)S(s_2)}ds_1ds_2 \\
&\ll x_1^3\sum_{m_1,m_2\in\mathcal{S}\cap[T,2T]}\int_{c+im_1}^{c+i(m_1+1)}\int_{c+im_2}^{c+i(m_2+1)}\frac{|C_1(s_1)S(s_1)C_1(s_2)S(s_2)|}{|1+s_1+\overline{s}_2|}ds_1ds_2 \\
&\ll x_1^{2\sigma+1}\tau^{-2}\sum_{m_1,m_2\in\mathcal{S}\cap[T,2T]}\int_{m_1}^{m_1+1}\int_{m_2}^{m_2+1}\frac{1}{1+|t_1-t_2|}dt_1dt_2 \\
&\ll \frac{x_1^{2\sigma+1}\log x_1}{\tau^2}\sum_{m_1\in\mathcal{S}\cap[T,2T]}1 \\
&\ll x_1^{1+2\sigma+\epsilon}\tau^{-2}R(T). \qquad (95)
\end{aligned}$$

(iii): L^4 estimate.

$$\begin{aligned}
&\int_x^{2x}|E((N_i),(\sigma_i);y,T)|^4dy \\
&\ll \int_x^{2x}\left|\sum_{m\in\mathcal{S}\cap[T,2T]}\int_m^{m+1}y^{c+it}C_1(c+it)S(c+it)dt\right|^4dy \\
&\ll \sum_{m_1,m_2,m_3,m_4\in\mathcal{S}\cap[T,2T]}\int_{m_1}^{m_1+1}\int_{m_2}^{m_2+1}\int_{m_3}^{m_3+1}\int_{m_4}^{m_4+1} \\
&\qquad \left|\int_x^{2x}y^{4c+i(t_1+t_2-t_3-t_4)}dy\right|\prod_{j=1}^4\left(\left|C_1(c+it_j)S(c+it_j)\right|dt_j\right) \\
&\ll x_1^5\left(\tau^{-4}\left(\prod_i N_i^{-4c+4\sigma_i}\right)\sum_{m_j\in\mathcal{S}\cap[T,2T]}\frac{1}{1+|m_1+m_2-m_3-m_4|}\right) \\
&\ll \frac{x_1^{1+4\sigma}}{\tau^4}\sum_{m_j\in\mathcal{S}\cap[T,2T]}\frac{1}{1+|m_1+m_2-m_3-m_4|}. \qquad (96)
\end{aligned}$$

We wish to bound the inner sum. We let

$$g(v) := \#\{(m_1, m_2, m_3, m_4) \in (\mathcal{S} \cap [T, 2T])^4 : m_1 + m_2 - m_3 - m_4 = v\}, \tag{97}$$

$$\mathcal{S}^* = \{(m_1, m_2 m_3, m_4) \in \mathcal{S}^4 : m_1 + m_2 - m_3 - m_4 = 0\}. \tag{98}$$

Then

$$\sum \frac{1}{1 + |m_1 + m_2 - m_3 - m_4|} \ll \sum_{|v| \le 4T} \frac{g(v)}{1 + |v|}. \tag{99}$$

But we have

$$\begin{aligned} g(v) &= \int_0^1 \Bigg| \sum_{m \in \mathcal{S} \cap [T, 2T]} e(mu) \Bigg|^4 e(-vu) du \\ &\le \int_0^1 \Bigg| \sum_{m \in \mathcal{S} \cap [T, 2T]} e(mu) \Bigg|^4 du = g(0). \end{aligned} \tag{100}$$

Hence

$$\sum_{m_j \in \mathcal{S} \cap [T, 2T]} \frac{1}{1 + |m_1 + m_2 - m_3 - m_4|} \ll \#\left(\mathcal{S}^* \cap [T, 2T]\right) \log T. \tag{101}$$

This gives us

$$\int_x^{2x} |E((N_i), (\sigma_i); y, T)|^4 dy \ll x_1^{1+4\sigma+\epsilon} \tau^{-4} R^*(T). \tag{102}$$

$\square$

4.3. **Estimation of $\sum d_n^2$.** We now use Lemma 4.2 to estimate $\sum d_n^2$.

Suppose $p_{n+1} - p_n \ge 4x/\tau$ and $x \le p_n \le 2x$. Let

$$y \in (p_n, (p_{n+1} + p_n)/2) \tag{103}$$

so that, for $x \le y \le 2x$, we have

$$p_n < y < y + y/\tau \le p_{n+1}. \tag{104}$$

Hence there are no primes in the interval $(y, y + y/\tau)$. In this case, by (69) we have

$$\psi(y + y/\tau) - \psi(y) = o\left(\frac{y}{\tau}\right). \tag{105}$$

Thus Lemma 4.2 yields

$$\sup_{(N_i),(\sigma_i),T} E((N_i), (\sigma_i); y, T) + E_6 \gg \frac{x}{\tau(\log x)^{4k+1}}. \tag{106}$$

We now wish to show that this cannot be the case too frequently.

Since $E_6 = 0$ for $y \notin \mathcal{B}_{y^{1/3}}$, we see that

$$E_6 \gg \frac{x}{\tau(\log x)^{4k+1}} \tag{107}$$

can only hold on a set of measure at most

$$\mathrm{meas}(\mathcal{B}_{y^{1/3}}) \ll x^{43/60+\epsilon} \tag{108}$$

by (33) and (17).

Suppose that for some choice of $(N_i), (\sigma_i), T$ we have

$$E((N_i), (\sigma_i); y, T) \gg \frac{x}{\tau(\log x)^{4k+1}}. \tag{109}$$

By (84) we must have

$$R(T) \gg x_1^{1-\sigma}(\log x_1)^{-4k-1}. \tag{110}$$

We now wish to estimate how frequently (109) can occur. By (33) and (85) we see that (109) can hold on a set of measure

$$\ll x_1^{2\sigma-1+\epsilon}R(T). \tag{111}$$

Similarly from (86) we see that this set has measure

$$\ll x_1^{4\sigma-3+\epsilon}R^*(T). \tag{112}$$

Therefore (109) holds on a set of measure

$$\ll \begin{cases} \min\left(x_1^{2\sigma-1+\epsilon}R(T), x_1^{4\sigma-3+\epsilon}R^*(T)\right), & R(T) \gg x_1^{1-\sigma}(\log x_1)^{-4k-1} \\ 0, & \text{otherwise.} \end{cases} \tag{113}$$

There are $O(x_1^{\epsilon})$ choices for $(N_i)_1^{2k}$, $(\sigma_1)_1^{2k}$ and T. Therefore

$$\sup_{(N_i),(\sigma_i),T} E((N_i), (\sigma_i); y, T) \gg \frac{x}{\tau(\log x)^{4k+1}} \tag{114}$$

can only hold on a set of cardinality

$$\ll x_1^{\epsilon} \sup_{\substack{(N_i),(\sigma_i),T \\ T\in[T_1,T_0] \\ R(T)\gg x_1^{1-\sigma}(\log x_1)^{-4k-1}}} \left(x_1^{\epsilon}\min\left(x_1^{2\sigma-1}R(T), x_1^{4\sigma-3}R^*(T)\right)\right). \tag{115}$$

Putting (108) and (115) together, we see that the set of y such that $y \in (p_n, p_n/2 + p_{n+1}/2)$ with $p_n+1 - p_n \geq 4x/\tau$ and $x \leq p_n \leq 2x$ must have measure

$$\ll \sup_{\substack{(N_i),(\sigma_i),T \\ T\in[T_1,T_0] \\ R(T)\gg x_1^{1-\sigma}(\log x_1)^{-4k-1}}} \left(x_1^{2\epsilon}\min\left(x_1^{2\sigma-1}R(T), x_1^{4\sigma-3}R^*(T)\right)\right) + x_1^{43/60+\epsilon}. \tag{116}$$

However, this set trivially has measure

$$\geq \sum_{\substack{p_{n+1}-p_n\geq 4x/\tau \\ p_n\geq x \\ (p_n+p_{n+1})/2\leq 2x}} \frac{p_{n+1}-p_n}{2}. \tag{117}$$

Therefore we have

$$\sum_{\substack{p_{n+1}-p_n\geq 4x/\tau \\ p_n\geq x \\ (p_n+p_{n+1})/2\leq 2x}} \frac{p_{n+1}-p_n}{2} \ll \sup_{\substack{(N_i),(\sigma_i),T \\ T\in[T_1,T_0] \\ R(T)\gg x_1^{1-\sigma}(\log x_1)^{-4k-1}}} \left(x_1^{2\epsilon}\min\left(x_1^{2\sigma-1}R(T), x_1^{4\sigma-3}R^*(T)\right)\right) + x_1^{43/60+\epsilon}. \tag{118}$$

There is at most one prime p_n such that $p_n \leq 2x < (p_n + p_{n+1})/2$. Hence

$$\sum_{\substack{4x/\tau\leq p_{n+1}-p_n\leq 8x/\tau \\ x\leq p_n\leq 2x}} (p_{n+1}-p_n) \ll \sup_{\substack{(N_i),(\sigma_i),T \\ T\in[T_1,T_0] \\ R(T)\gg x_1^{1-\sigma}(\log x_1)^{-4k-1}}} \left(x_1^{2\epsilon}\min\left(x_1^{2\sigma-1}R(T), x_1^{4\sigma-3}R^*(T)\right)\right) + x_1^{43/60+\epsilon} + \frac{x_1}{\tau}. \tag{119}$$

Thus

$$\sum_{\substack{4x/\tau\le p_{n+1}-p_n\le 8x/\tau\\ x\le p_n\le 2x}}(p_{n+1}-p_n)^2 \ll \sup_{\substack{(N_i),(\sigma_i),T\\ T\in[T_1,T_0]\\ R(T)\gg x_1^{1-\sigma}(\log x_1)^{-4k-1}}}\Big(x_1^{2\epsilon}\tau^{-1}\min\Big(x_1^{2\sigma}R(T),x_1^{4\sigma-2}R^*(T)\Big)\Big)$$
$$+\frac{x_1^{103/60+\epsilon}}{\tau}+\frac{x_1^2}{\tau^2}. \tag{120}$$

This is good enough to prove that

$$\sum_{\substack{4x/\tau\le d_n\le 8x/\tau\\ x\le p_n\le 2x}} d_n^2\ll x^{1+\nu+10\epsilon} \tag{121}$$

if we can prove that

$$\sup_{\substack{(N_i),(\sigma_i),T\\ T\in[T_1,T_0]\\ R(T)\gg x_1^{1-\sigma}(\log x_1)^{-4k-1}}}\Big(x_1^{2\epsilon}\min\Big(x_1^{2\sigma}\tau^{-1}R(T),x_1^{4\sigma-2}\tau^{-1}R^*(T)\Big)\Big)+x_1^{103/60+\epsilon}/\tau\ll x_1^{1+\nu+10\epsilon}. \tag{122}$$

Therefore (recalling $T_0=\tau(\log x)^3$) we have proven the following proposition.

Proposition 4.3. *Let* $x^{19/40}\le\tau\le x^{3/4-\epsilon}$ *and* $\nu\ge 29/120$.

If, uniformly for all $T\in[T_1,T_0]$ *and for all possible* $(N_i),(\sigma_i)$ *satisfying* (76), (77), (78) *and* (79), *at least one of the following holds:*

$$(i): R(T)\ll x_1^{1-\sigma}(\log x_1)^{-4k-2}, \tag{123}$$

$$(ii): R(T)\ll T_0x_1^{1+\nu-2\sigma+8\epsilon}, \tag{124}$$

$$(iii): R^*(T)\ll T_0x_1^{3+\nu-4\sigma+8\epsilon}, \tag{125}$$

then we have

$$\sum_{\substack{4x/\tau\le d_n\le 8x/\tau\\ x\le p_n\le 2x}} d_n^2\ll x^{1+\nu+10\epsilon}. \tag{126}$$

5. Large Values of Dirichlet Polynomials

We recall that

$$x_1 = \prod_{i=1}^{2k} N_i, \quad x_1^{\sigma} = \prod_{i=1}^{2k} N_i^{\sigma_i}, \quad N_i \ll x_1^{19/20} \quad \forall i,$$

$$N_i \ll x_1^{1/k} \quad \text{if } i \leq k, \quad \sigma_i \leq 1 \quad \forall i. \tag{127}$$

In this section we aim to use published large value estimates to obtain bounds on $R(T)$ and $R^*(T)$. Specifically we aim to prove the following proposition.

Proposition 5.1. *One of the following holds uniformly for $T \in [T_1, T_0]$ and for any $(N_i), (\sigma_i)$ satisfying* (127)

$$\begin{aligned} (i): &\quad R(T) \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}, \\ (ii): &\quad R(T) \ll T_0 x_1^{5/4-2\sigma+8\epsilon}, \\ (iii): &\quad R^*(T) \ll T_0 x_1^{13/4-4\sigma+8\epsilon}. \end{aligned}$$

Hence Proposition 3.2 holds by Proposition 4.3 (with $\nu = 1/4$) and (17).

Heath-Brown used essentially the same argument thus far in [11] and [12], but worked with zeroes of $\zeta(s)$ instead of Dirichlet polynomials. The estimates on the density of zeroes used the zero detection method, which amounted to bounding the frequency with which Dirichlet polynomials take large values. The advantage we get from using Dirichlet polynomials throughout is that we have the additional condition that the total combined length x_1 of $S_1 \ldots S_{2k}$ is approximately x. If we did not have this restriction to our Dirichlet polynomials then we would only be able to produce the same result as Heath-Brown [12] . The critical case would have been when $\sigma_i = 3/4\ \forall i$, $x_1 = \tau^{9/5}$ and $N_i = \tau^{2/5}$ or $1/2\ \forall i$. But we cannot have a set of Dirichlet polynomials each with length $\tau^{2/5}$ and combined length $\tau^{9/5}$, and so the critical case cannot exist when we have this additional constraint. This allows us to improve the overall result.

We put

$$\mu = \frac{\log x_1}{\log T_0} \tag{128}$$

to simplify notation. we note that since $x \ll x_1 \ll x$, inequality (17) implies that we only need consider

$$\frac{4}{3} \leq \mu \leq \frac{19}{9}. \tag{129}$$

5.1. Initial Estimates.

Our proof will make extensive use of the following three bounds on the frequency of large values taken by Dirichlet polynomials.

We consider a Dirichlet polynomial $S(t) = \sum_N^{2N} a_n n^{-c+it}$ which is a product of some of the S_i. Therefore $S = \prod_{i\in\mathcal{I}} S_i$, $N = \prod_{i\in\mathcal{I}} N_i$ and $N^{\sigma'} = \prod_{i\in\mathcal{I}} N_i^{\sigma_i}$ for some set $\mathcal{I} \subset \{1, \ldots, 2l\}$.

We let

$$R(S;T)=\#\Big\{m\in\mathbb{Z}\cap[T,2T]:N^{-c+\sigma'}\le\sup_{t\in[m,m+1]}|S(t)|\le 2N^{-c+\sigma'}\Big\},$$

$$R^*(S;T)=\#\Big\{(m_1,m_2,m_3,m_4)\in(\mathbb{Z}\cap[T,2T])^4:m_1+m_2=m_3+m_4,$$
$$N^{-c+\sigma'}\le\sup_{t\in[m_i,m_i+1]}|S(t)|\le 2N^{-c+\sigma'}\forall i\Big\}.$$

Clearly we have $R(T)\le R(S;T)$ and $R^*(T)\le R^*(S;T)$. We note that the coefficients a_n of S satisfy $a_n=O_\delta(T_0^\delta)$ for every $\delta>0$.

Lemma 5.2 (Montgomery's Mean Value Estimate)**.** *We have*

$$R(S;T)\ll(\log NT)\left(N^{2-2\sigma'}+TN^{1-2\sigma'}\right)\left(\frac{\sum_N^{2N}|a_n|^2}{N}\right).$$

In particular, uniformly for $T_1\le T\le T_0$ and for any $\delta>0$, we have

$$R(T)\ll_\delta T_0^\delta N^{2-2\sigma'}+T_0^{1+\delta}N^{1-2\sigma'}.$$

Proof. The first statement is proven in [22][Theorem 7.3]. The second statement follows immediately from the first since $N\le x_1\le T_0^3$. □

Lemma 5.3 (Huxley's Large Values Estimate)**.** *We have*

$$R(S;T)\ll(\log NT)^2\left(N^{2-2\sigma}+TN^{4-6\sigma}\right)\left(1+\frac{\sum_N^{2N}|a_n|^2}{N}\right)^3.$$

In particular, uniformly for $T_1\le T\le T_0$ and for any $\delta>0$, we have

$$R(T)\ll_\delta T_0^\delta N^{2-2\sigma'}+T_0^{1+\delta}N^{4-6\sigma'}.$$

Proof. The first statement is proven in [18][Equation 2.9]. The second statement follows immediately from the first since $N\le x_1\le T_0^3$. □

Lemma 5.4 (Heath-Brown's R^* Bound)**.** *For any $\delta>0$ we have*

$$R^*(S;T)\ll_\delta N^{1-2\sigma'}T^\delta(R(S;T)N+R(S;T)^2+R(S;T)^{5/4}T^{1/2})^{1/2}$$
$$\times(R^*(S;T)N+R(S;T)^4+R(S;T)R^*(S;T)^{3/4}T^{1/2})^{1/2}.$$

In particular, uniformly for $T_1\le T\le T_0$ and for any $\delta>0$, we have

$$R^*(T)\ll_\delta N^{1-2\sigma'}T_0^\delta(R(T)N+R(T)^2+R(T)^{5/4}T_0^{1/2})^{1/2}$$
$$\times(R^*(T)N+R(T)^4+R(T)R^*(T)^{3/4}T_0^{1/2})^{1/2}$$

Proof. The first statement is proven in [13][Equation 33]. The second statement follows from a precisely analogous argument applied to $R^*(T)$. □

In addition to these results, we will also require the following lemma.

Lemma 5.5. *Either*

$$R(T)\ll x^{1-\sigma}(\log x)^{-4k-2}$$

or for $k<i\le 2k$ we have

$$R(T)\ll(\log T)^{41}T^2N_i^{6-12\sigma_i}\quad\textit{and}\quad R(T)\ll(\log T)^{13}TN_i^{2-4\sigma_i}.$$

Proof. We follow the method of Heath-Brown in [14] but making use of the twelfth as well as the fourth power moment of the Zeta function.

We consider a polynomial S_i with $k < i \leq 2k$. Such a polynomial has all coefficients 1 (if $i < 2k$) or all coefficients $\log n$ (if $i = 2k$). We first consider the case when the coefficients of S_i are identically 1.

From Perron's formula with $T_1 \leq T \leq T_0$ we have for $T \leq t \leq 2T$ that

$$\left|\sum_{N}^{2N} n^{-1/2-it}\right| = \left|\int_{d-iT/2}^{d+iT/2} \zeta(1/2+it+s)\frac{(2N)^s - N^s}{s}ds\right| + O(N^{1/2}T^{-1}(\log x_1) + 1)$$

where $d = 1/2 + (\log x_1)^{-1}$.

Moving the line of integration to $\Re(s) = 0$ gives

$$\left|\sum_{N}^{2N} n^{-1/2-it}\right| \ll \int_{T/2}^{5T/2} |\zeta(1/2+iu)| \frac{du}{1+|t-u|} + N^{1/2}T^{-1}(\log x_1) + 1.$$

Let $(m_j)_1^{R(T)} \subset \mathbb{Z} \cap [T, 2T]$ be such that

$$\sup_{t\in[m_j,m_j+1]} \left|\sum_{N}^{2N} n^{-1/2-it}\right| \gg N^{\sigma-1/2}.$$

Let t_j be a point in $[m_j, m_j+1]$ where this supremum is attained.

Then, using Hölder's inequality and Heath-Brown's twelfth power moment bound for $\zeta(s)$ (see [9][Theorem 1]):

$$\begin{aligned}
R(T)N^{12\sigma-6} &\ll \sum_{1\leq j\leq R(T)} \left|\sum_{N}^{2N} n^{-1/2+it_j}\right|^{12} \\
&\ll \sum_{1\leq j\leq R(T)} \left(\int_T^{2T} |\zeta(1/2+iu)|^{12} \frac{du}{1+|t_j-u|}\right)\left(\int_T^{2T} \frac{du}{1+|t_j-u|}\right)^{11} \\
&\qquad + R(T)N^6T^{-12}(\log x_1)^{12} + R(T) \\
&\ll (\log x_1)^{11} \int_T^{2T} |\zeta(1/2+iu)|^{12} \sum_{1\leq j\leq R(T)} \frac{1}{1+|t_j-u|}du \\
&\qquad + R(T)N^6T^{-12}(\log x_1)^{12} + R(T) \\
&\ll T^2(\log x_1)^{29} + R(T)N^6T^{-12}(\log x_1)^{12} + R(T) \\
&\ll T^2(\log x_1)^{29} + R(T)N^6T^{-12}(\log x_1)^{12},
\end{aligned}$$

since $R(T) \ll T$.

In the case $i = 2k$ and all coefficients are $\log n$ we obtain by partial summation and the method above

$$R(T)N^{12\sigma-6} \ll T^2(\log x_1)^{41} + R(T)N^6T^{-12}(\log x_1)^{24}.$$

In either case we get

$$R(T) \ll (\log x_1)^{41}\left(T^2N^{6-12\sigma} + R(T)N^{12-12\sigma}T^{-12}\right). \tag{130}$$

We can apply the same method, but using the fourth power moment of $\zeta(s)$ (see [19][Theorem B] for example) instead of the twelfth. We obtain (for both $\sum_N^{2N} n^{-1/2-it}$ and $\sum_N^{2N}(\log n)n^{-1/2-it}$)

$$\begin{aligned} R(T)N^{4\sigma-2} &\ll (\log x_1)^8\left(\int_{T/2}^{5T/2}|\zeta(1/2+iu)|^4\sum_{1\le j\le R(T)}\frac{1}{1+|t_j-u|}du\right)\\ &\quad+(\log x_1)^8\left(R(T)N^2T^{-6}+R(T)\right)\\ &\ll(\log x_1)^{13}\left(T+R(T)N^2T^{-4}\right).\end{aligned}$$

Thus

$$R(T)\ll(\log x_1)^{13}\left(TN^{2-4\sigma}+R(T)N^{4-4\sigma}T^{-4}\right). \tag{131}$$

From (130) and (131) we see one of the following must hold for any Dirichlet polynomial S_j with $i>k$:

(i): $T\ll(\log x_1)^4N_j^{1-\sigma_j}$

(ii): $R(T)\ll(\log x_1)^{41}T^2N_j^{-6(2\sigma_j-1)}$ and $R(T)\ll(\log x_1)^{13}TN_j^{2-4\sigma_j}$.

We are therefore left to show that (i) implies that $R(T)\ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$.

We note that

$$R(T)\ll T\ll(\log x_1)^4N_j^{1-\sigma_j}\ll(\log x_1)^4x_1^{1-\sigma}\prod_{i\neq j}N_i^{\sigma_i-1}.$$

This is good enough to prove

$$R(T)\ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$$

provided that for some $j'\neq j$ we have

$$N_{j'}^{1-\sigma_{j'}}\gg(\log x_1)^{4k+6}.$$

Since $N_i\ll x_1^{19/20}$ $\forall i$ there must be some $j'\neq j$ such that $N_{j'}\gg x_1^{1/40k}$ (since there are $2k$ polynomials whose combined length $\prod N_i$ is x_1). Thus we need to show that $\sigma_{j'}$ cannot be too close to 1.

We put

$$\eta=\eta(x_1)=C_0(\log x_1)^{-2/3}(\log\log x_1)^{-1/3} \tag{132}$$

for some suitable constant $C_0>0$ (which we will declare later).

By Perron's formula we have for $t\in[T,2T]$ that

$$\left|S_{j'}(c+it)\right|=\frac{1}{2\pi i}\int_{-iT/2}^{iT/2}F_{j'}(c+it+s)\frac{(2N_{j'})^s-N_{j'}^s}{s}ds+O(T^{-1}\log x_1)$$

where

$$F_{j'}(s)=\begin{cases}(\zeta(s))^{-1}, & 1\le j'\le k\\ \zeta(s), & k<j'<2k\\ \zeta'(s), & j'=2k\end{cases}$$

In the region $1-2\eta-c\le\Re(s)\le 0$, $|t-\Im(s)|\le T/2$ we have

$$\left|F_{j'}(c+it+s)\right|\ll(\log x_1)^2$$

for any $1\le j'\le 2k$. This follows from [30][Theorem 3.11] along with the Vinogradov-Korobov estimate as given in [26] for a suitable choice of C_0.

We now move the line of integration to $\Re(s) = 1 - 2\eta - c$ to obtain

$$|S_{j'}(c+it)| \ll \int_{T/2}^{5T/2} |F_j(1-2\eta+it+is)| N_{j'}^{-2\eta} |s|^{-1} ds + O(T^{-1}\log x_1)$$
$$\ll (\log x_1)^3 (N_{j'}^{-2\eta} + T_1^{-1}).$$

Thus, since $N_{j'} > x_1^{1/40k}$, we have $N_{j'}^{\eta/2} \gg (\log x_1)^4$. This gives

$$|S_{j'}(c+it)| \leq N_i^{-3\eta/2}.$$

Therefore we have

$$R = 0 \qquad \text{or} \qquad \sigma_{j'} \leq 1 - 3\eta/2 \tag{133}$$

for any polynomial with $N_{j'} > x_1^{1/40k}$.

In particular either

$$R = 0 \ll x_1^{1-\sigma} (\log x)^{-4k-2}$$

or

$$N_{j'}^{1-\sigma_{j'}} \gg \left(x_1^{1/40k}\right)^{3\eta/2} \gg (\log x_1)^{4k+6}$$

which implies that

$$R(T) \ll x_1^{1-\sigma} (\log x_1)^{-4k-6}.$$

Thus the lemma holds. □

We now summarise the other large-value estimates which we will make use of, which follow from published work by other authors.

Lemma 5.6. *Either:*

$$R(T) \ll x_1^{1-\sigma} (\log x_1)^{-4k-2}$$

or:
uniformly for $T_1 \leq T \leq T_0$ we have

$$R(T) \ll_\delta T_0^{(3-3\sigma)/(2-\sigma)+\delta}, \qquad \text{if } \sigma \leq 3/4, \tag{134}$$

$$R(T) \ll_\delta T_0^{(3-3\sigma)/(3\sigma-1)+\delta}, \qquad \text{if } \sigma \geq 3/4, \tag{135}$$

$$R(T) \ll_\delta T_0^{(3-3\sigma)/(10\sigma-7)+\delta}, \qquad \text{if } \sigma \leq 25/28, \tag{136}$$

$$R(T) \ll_\delta T_0^{(4-4\sigma)/(4\sigma-1)+\delta}, \qquad \text{if } \sigma \geq 25/28, \tag{137}$$

$$R^*(T) \ll_\delta T_0^{(15-16\sigma)/2+\delta}, \qquad \text{if } \sigma \leq 3/4, \tag{138}$$

$$R^*(T) \ll_\delta T_0^{(12-12\sigma)/(4\sigma-1)+\delta}, \qquad \text{if } \sigma \geq 3/4 \tag{139}$$

for any $\delta > 0$.

Proof. We assume that $R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$ does not hold.

These bounds are usually obtained merely as an intermediate step in the zero detection method when trying to bound $N(\sigma, T)$ (or $N^*(\sigma, T)$). They are therefore not always explicitly stated as a lemma in the papers where they are obtained.

Since the published bounds all bound Dirichlet polynomials which arise from the zero detection method, they do not immediately apply to our context, since the Dirichlet polynomials we are considering can be slightly different. In particular, the results we will

quote only apply to a Dirichlet polynomial S with length $N \in [Y^{1/2}, Y]$ (for some value of $Y \le T_0$) and coefficients which are $O_\delta(T_0^\delta)$ for every $\delta > 0$.

We repeatedly combine any pair of polynomials of length $\le T_0^\delta$, so that there is at most one polynomial of length $\le T_0^\delta$. We only need to consider $\delta < 1/(2k)$, and so any polynomial with length $\ge T_0^{3/k}$ must have all coefficients 1 all coefficients $\log n$. This means that the R-bounds of Lemma 5.5 still apply to any of the polynomials with length $\ge T_0^{3/k}$ after these combinations.

We pick a polynomial S_{j_1} of length $N_{j_1} > T_0^\delta$ with σ_{j_1} maximal.

If $\sigma_{j_1} < \sigma$ then, since σ is an average of the σ_i, the polynomial S_{j_2} with length $N_{j_2} \le T_0^\delta$ must exist and have $\sigma_{j_2} > \sigma$. In this case we combine the polynomials S_{j_1} and S_{j_2} to produce a polynomial S of length N and size $\sigma' \ge \sigma$.

If $\sigma_{j_1} \ge \sigma$ we take $S = S_{j_1}$ (and so $N = N_{j_1}$ and $\sigma' = \sigma$).

Since the bounds (134), (135), (136), (137), (138) and , (139) are all decreasing in σ, it is sufficient to prove them for $R(S;T)$ and $R^*(S;T)$ when $\sigma' = \sigma$.

If $N \le Y$, then by raising the polynomial S to a suitable exponent we can ensure that the new polynomial, S' say, has length $N' \in [Y^{1/2}, Y]$. The Dirichlet polynomials S_i which we are considering (or any combination of them) have coefficients which are $O_\delta(T_0^\delta)$ for every $\delta > 0$. Therefore the coefficients of S' will also be $O_\delta(T_0^\delta)$ for every $\delta > 0$ provided we have raised S to an exponent which is $O_\delta(1)$. This is the case since by construction we have $N > T_0^\delta$. Therefore the published bound will hold if $N \le Y$.

If $N \ge Y$ then we will use Lemma 5.5 to obtain the result (recalling that we have assumed that $R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$ does not hold). If $N = N_{j_1}$ or $N = N_{j_1}N_{J_2}$ then by choosing k large enough we must have $N_{j_1} > T_0^{3/k}$, and so Lemma 5.5 applies. If $N = N_{j_1}N_{j_2}$ then we have

$$R(T) \ll R(T)T_0^{2\delta}N_{j_2}^{2-4\sigma_{j_2}} \ll_\delta T_0^{1+3\delta}N_{j_1}^{2-4\sigma_{j_1}}N_{j_2}^{2-4\sigma_{j_2}} \ll_\delta T_0^{1+3\delta}N^{2-4\sigma} \tag{140}$$

and

$$R(T) \ll R(T)T_0^{6\delta}N_{j_2}^{6-12\sigma_{j_2}} \ll_\delta T_0^{1+7\delta}N^{6-12\sigma} \tag{141}$$

for any $\delta > 0$. We see that (140) and (141) trivially follow from Lemma 5.5 if $N = N_{j_1}$, and so they hold in either case.

We now establish (134), (135), (136), (137), (138) and (139) in turn.

We see that (134) holds trivially if $\sigma \le 2/3$. In the proof of Theorem 12.1 in [22], Montgomery shows that $R(S;T) \ll_\delta T_0^{(3-3\sigma)/(2-\sigma)+\delta}$ if S has length $N \in [Y^{1/2}, Y]$ with $Y = T_0^{3/(8-4\sigma)}$ and $1/2 \le \sigma \le 3/4$. Therefore (134) holds if $N \le T_0^{3/(8-4\sigma)}$ (since $N > T_0^\delta$). If $N \ge T_0^{3/(8-4\sigma)}$ then since $k \ge 9$ we must have $N_{j_1} > T_0^{3/k}$. Then by (140) we have

$$R(T) \ll_\delta T_0^{1+3\delta}N^{2-4\sigma} \ll T_0^{(7-8\sigma)/(4-2\sigma)+3\delta} \ll T_0^{(3-3\sigma)/(2-\sigma)+3\delta}$$

for any $\delta > 0$ (since we only need to consider $\sigma \ge 2/3$). This establishes (134).

In the proof of inequality (19) in [18], Huxley shows that $R(S;T) \ll_\delta T_0^{(3-3\sigma)/(3\sigma-1)+\delta}$ if S has length $N \in [Y^{1/2}, Y]$ with $Y = T_0^{3/(12\sigma-4)}$. Therefore (135) holds if $N \le T_0^{3/(12\sigma-4)}$

(since $N > T_0^{\delta}$). If $N \geq T_0^{3/(12\sigma-4)}$ then since $k \geq 9$ we must have $N_{j_1} > T_0^{3/k}$. Then by (141) we have

$$R(T) \ll_\delta T_0^{2+7\delta} N^{6-12\sigma} \ll T_0^{(5-6\sigma)/(6\sigma-2)+7\delta} \ll T_0^{(3-3\sigma)/(3\sigma-1)+7\delta}$$

for any $\delta > 0$. This establishes (135).

In the proof of Theorem 1 in [13], Heath-Brown proves $R(S\,;T) \ll_\delta T_0^{(3-3\sigma)/(10\sigma-7)+\delta}$ if S has length $N \in [Y^{1/2}, Y]$ with $Y = T_0^{3/(40\sigma-28)}$ and $\sigma \leq 25/28$. Therefore (136) holds if $N \leq T_0^{3/(40\sigma-28)}$ (since $N > T_0^{\delta}$). If $N \geq T_0^{3/(40\sigma-28)}$ then since $k \geq 13$ we must have $N_{j_1} > T_0^{3/k}$. Then by (141) we have

$$R(T) \ll_\delta T_0^{2+7\delta} N^{6-12\sigma} \ll T_0^{(22\sigma-19)/(20\sigma-14)+7\delta} \ll T_0^{(3-3\sigma)/(10\sigma-7)+7\delta}$$

for any $\delta > 0$ (since we are only considering $\sigma \leq 25/28$ in (136)). This establishes (136).

In the proof of Theorem 1 in [13], Heath-Brown shows that $R(S\,;T) \ll_\delta T_0^{(4-4\sigma)/(4\sigma-1)+\delta}$ if S has length $N \in [Y^{1/2}, Y]$ with $Y = T_0^{1/(4\sigma-1)}$ and $\sigma \geq 25/28$. Therefore (137) holds if $N \leq T_0^{1/(4\sigma-1)}$ (since $N > T_0^{\delta}$). If $N \geq T_0^{1/(4\sigma-1)}$ then since $k \geq 10$ we must have $N_{j_1} > T_0^{3/k}$. Then by (141) we have

$$R(T) \ll_\delta T_0^{2+7\delta} N^{6-12\sigma} \ll T_0^{(4-4\sigma)/(4\sigma-1)+7\delta}$$

for any $\delta > 0$. This establishes (137).

We see that (138) holds trivially if $\sigma \leq 1/2$. In the proof of Theorem 2 in [13], Heath-Brown shows that $R^*(S\,;T) \ll_\delta T_0^{(10-11\sigma)/(2-\sigma)+\delta} + T_0^{(18-19\sigma)/(4-2\sigma)+\delta}$ if S has length $N \in [Y^{1/2}, Y]$ with $Y = T_0^{1/2}$ and $1/2 \leq \sigma \leq 3/4$. In particular, this gives $R(S\,;T) \ll_\delta T_0^{(15-16\sigma)/2+\delta}$ for any $\delta > 0$ and $\sigma \leq 3/4$. Therefore (138) holds if $N \leq T_0^{1/2}$ (since $N > T_0^{\delta}$). If $N \geq T_0^{1/2}$ then since $k \geq 7$ we must have $N_{j_1} > T_0^{3/k}$. In this case, using the trivial bound $R^*(T) \ll (\log T_0) R(T)^3$ and (140), we have

$$R^*(T) \ll_\delta (\log T_0)(T_0^{1+3\delta/4} N^{2-4\sigma})^3 \ll T_0^{6-6\sigma+10\delta} \ll T_0^{(15-16\sigma)/2+10\delta}$$

for any $\delta > 0$ (since we are only considering $\sigma \leq 3/4$ in (138)). This establishes (138).

In the proof of Theorem 2 in [13], Heath-Brown proves $R^*(S\,;T) \ll_\delta T_0^{(12-12\sigma)/(4\sigma-1)+\delta}$ if S has length $N \in [Y^{1/2}, Y]$ with $Y = T_0^{1/(4\sigma-1)}$ and $\sigma \geq 3/4$. Therefore (139) holds if $N \leq T_0^{1/(4\sigma-1)}$ (since $N > T_0^{\delta}$). If $N \geq T_0^{1/(4\sigma-1)}$ then since $k \geq 10$ we must have $N_{j_1} > T_0^{3/k}$. Then by (141) we have

$$R^*(T) \ll (\log T_0) R(T)^3 \ll_\delta (\log T_0)(T_0^{2+7\delta} N^{6-12\sigma})^3 \ll T_0^{(12-12\sigma)/(4\sigma-1)+22\delta}$$

for any $\delta > 0$. This establishes (139). □

To simplify notation we drop the T from R and R^* since we are from now on only interested in them evaluated at T. Thus

$$R = R(T), \qquad R^* = R^*(T).$$

We now prove Proposition 5.1 by way of five lemmas. Lemma 5.7 covers the case when some of the polynomials are long by using Lemma 5.5. Lemma 5.8 covers the case $\sigma \leq 3/4$ by using Montgomery's mean-value estimate and Heath-Browns R^* estimate. Lemma 5.9 covers the case $\sigma \geq 3/4$ and μ 'small' using the same method but using Huxley's large values estimate and Heath-Brown's R^* estimate. Lemma 5.10 covers the case when

$\sigma > 3/4$ and μ is 'large' using an adapted argument from [13] and Lemma 5.5. Lemma 5.11 deals with the range when σ is very close to 1 using Vinogradov's zero-free region of $\zeta(s)$ and Van-der-Corput's method of exponential sums.

5.2. **Part 1: Long Polynomials.** We first notice that we only need to consider polynomials of reasonably short length, where published estimates for the frequency with which they take large values apply.

Lemma 5.7. *Either we have one of*

$$R \ll T_0 x_1^{5/4-2\sigma+\epsilon}, \quad R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$$

or we have

$$N_i \leq T_0^{1/2+\epsilon} \tag{142}$$

for all but at most one i. If such an exceptional polynomial S_j exists then $T_0^{1/2+\epsilon} \leq N_j \leq x_1^{3/5}$.

Proof. We assume that $R \ll x_1^{1-\sigma}(\log x)^{-4k-2}$ does not hold. Therefore the results of Lemmas 5.5 and 5.6 apply.

We consider polynomials S_i with $N_i > T_0^{1/2}$. Since we are taking $k \geq 6$, by (23) any such polynomial with 'long' length must be one where all coefficients are 1 or $\log(n)$. This means we can use Lemma 5.5 to get stronger than normal bounds.

Case 1: There are at least 2 such values of j such that $N_j > T_0^{1/2+\epsilon}$.

Let j_1, j_2 be two values of j such that $N_j > T_0^{1/2+\epsilon}$. We let $N = N_{j_1}N_{j_2}$ $(> T_0^{1+2\epsilon})$, $N^\alpha = N_{j_1}^{\sigma_{j_1}}N_{j_2}^{\sigma_{j_2}}$, $M = \prod_{i\neq j_1,j_2} N_i$, $M^\beta = \prod_{i\neq j_1,j_2} N_i^{\sigma_i}$. By Lemmas 5.2, 5.3 and 5.5 we have for any $\delta > 0$ that

$$\begin{aligned}
R &\ll_\delta (T_0^{1+\delta}N_{j_1}^{2-4\sigma_{j_1}})^{1/2}(T_0^{1+\delta}N_{j_2}^{2-4\sigma_{j_2}})^{1/2} = T_0^{1+\delta}N^{1-2\alpha},\\
R &\ll_\delta (T_0^{2+\delta}N_{j_1}^{6-12\sigma_{j_1}})^{1/2}(T_0^{2+\delta}N_{j_2}^{6-12\sigma_{j_2}})^{1/2} = T_0^{2+\delta}N^{3-6\alpha},\\
R &\ll_\delta M^{2-2\beta}T_0^{\delta} + \min(T_0^{1+\delta}M^{1-2\beta}, T_0^{1+\delta}M^{4-6\beta}).
\end{aligned}$$

We choose $\delta = \epsilon/2$, and so the implied constants only need to depend on ϵ.

If $R \ll M^{2-2\beta}T_0^{\epsilon/2}$ then since $N \geq T^{1+2\epsilon}$ we have

$$\begin{aligned}
R &\ll (T_0^{1+\epsilon/2}N^{1-2\alpha})^{1/2}(T_0^{\epsilon/2}M^{2-2\beta})^{1/2}\\
&= T_0^{1/2+\epsilon/2}N^{-1/2}x_1^{1-\sigma}\\
&\ll x_1^{1-\sigma}T_0^{-\epsilon/2}\\
&\ll x_1^{1-\sigma}(\log x_1)^{-4k-2}.
\end{aligned}$$

If $R \ll \min(T_0^{1+\epsilon/2}M^{1-2\beta}, T_0^{1+\epsilon/2}M^{4-6\beta})$ then since $N \geq T$

$$\begin{aligned}
R &\ll (T_0^{1+\epsilon/2}M^{1-2\beta})^{1/4}(T_0^{1+\epsilon/2}M^{4-6\beta})^{1/4}(T_0^{1+\epsilon/2}N^{1-2\alpha})^{1/4}(T_0^{2+\epsilon/2}N^{3-6\alpha})^{1/4}\\
&= T_0^{5/4+\epsilon/2}N^{-1/4}x_1^{5/4-2\sigma}\\
&\ll T_0 x_1^{5/4-2\sigma}.
\end{aligned}$$

Therefore if there are two polynomials with length $\geq T_0^{1/2+\epsilon}$ then

$$R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2} \qquad \text{or} \qquad R \ll T_0 x_1^{5/4-2\sigma+\epsilon}.$$

Case 2: There is a j such that $N_j > x_1^{3/5}$.

We consider the long polynomial S_j and its complement. To ease notation we let $N = N_j$, $N^\alpha = N_j^{\sigma_j}$, $M = \prod_{i\neq j} N_i$ with $M^\beta = \prod_{i\neq j} N_i^{\sigma_i}$. Then by Lemmas 5.5 and 5.2 (choosing $\delta = \epsilon$) we have

$$\begin{aligned} R &\ll (M^{2-2\beta}T_0^{\epsilon} + T_0^{1+\epsilon}M^{1-2\beta})^{6/7}(T_0^{2+\epsilon}N^{6-12\alpha})^{1/7} \\ &\ll T_0^{2/7+\epsilon}x_1^{12(1-\sigma)/7}N^{-6/7} + T_0^{8/7+\epsilon}x_1^{6/7-12\sigma/7} \end{aligned}$$

Since $N > x_1^{3/5}$ and $4/3 \leq \mu \leq 19/9$ this gives

$$R \ll T_0 x_1^{5/4-2\sigma+\epsilon}.$$

Therefore the Lemma holds. □

We note that inequalities (23) and (142) are vital in our treatment of the problem in this way. The S_i for $i \leq k$ are 'difficult' since the coefficients $\mu(n)$ have complicated behaviour, but by increasing k we can ensure these polynomials do not cause too many problems. This is because we have effective bounds on the number of large values reasonably short Dirichlet polynomials can take. We do not have the same method of controlling the length of S_i for $i > k$, but these polynomials have 'well-behaved' coefficients. This allows us to produce much stronger bounds in Lemma 5.5 and so cope with the longer polynomials.

From now on we assume that $N_i \leq T_0^{1/2+\epsilon}$ $\forall i$ except for possibly one exceptional polynomial S_j with $T_0^{1/2+\epsilon} \leq N_j \leq x_1^{3/5}$.

5.3. **Part 2:** $\sigma \leq 3/4$.

Lemma 5.8. *Let $\sigma \leq 3/4$. Then either*

$$R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$$

or

$$R \ll T_0 x_1^{5/4-2\sigma+2\epsilon}$$

or

$$R^* \ll T_0 x_1^{13/4-4\sigma+6\epsilon}.$$

Proof. We assume that $R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$ does not hold. Therefore the results of Lemma 5.6 apply.

The result follows from published estimates of Lemma 5.6 unless $8/5 \leq \mu \leq 2$ and $7/10 \leq \sigma \leq 3/4$.

For $\sigma \leq 5/8$ we use the trivial estimate

$$R \ll T_0 \ll T_0 x^{5/4-2\sigma}.$$

By (134) (choosing $\delta = \epsilon$) we have

$$R \ll T_0^{(3-3\sigma)/(2-\sigma)+\epsilon}.$$

This gives $R \ll T_0 x^{5/4-2\sigma+\epsilon}$ if $\mu \leq 2$ and $5/8 \leq \sigma \leq 7/10$ or if $\mu \leq 8/5$ and $7/10 \leq \sigma \leq 3/4$.

By (138) (choosing $\delta = \epsilon$) we have

$$R^* \ll T_0^{(15-16\sigma)/2+\epsilon}.$$

This gives $R \ll T_0 x^{13/4-4\sigma+\epsilon}$ if $\mu \geq 2$.

These cover all ranges of μ and σ unless $8/5 \leq \mu \leq 2$ and $7/10 \leq \sigma \leq 3/4$. We now consider this case.

We combine the polynomials to produce two polynomials of length M, N and size α, β. (so $M = \prod_{i \in I_1} N_i$, $M^\alpha = \prod_{i \in I_1} N_i^{\sigma_i}$, $N = \prod_{i \in I_2} N_i$, $N^\beta = \prod_{i \in I_2} N_i^{\sigma_i}$ for some disjoint $I_1, I_2 \subset \{1, \dots 2k\}$ with $I_1 \cup I_2 = \{1, \dots, 2k\}$). We will declare how we combine the polynomials later. We let M be the smaller of the two (so $M \leq N$). Therefore we have

$$x_1 = T_0^\mu = MN, \qquad M^\alpha N^\beta = x_1^\sigma.$$

Since $\mu \leq 2$, we have $T_0^2 \geq x_1 = MN \geq M^2$, and so $M \leq T_0$.

By Lemma 5.2 (choosing $\delta = \epsilon$) we have

$$R \ll \min(T_0^\epsilon M^{2-2\alpha} + T_0^{1+\epsilon} M^{1-2\alpha}, T_0^\epsilon N^{2-2\beta} + T_0^{1+\epsilon} N^{1-2\beta}).$$

We note that the first term in each component dominates iff the polynomial has length $\geq T_0$. Since $M \leq T_0$, we always have the second term ($T_0^{1+\epsilon} M^{1-2\alpha}$) dominating the first component of the minimum. We split the argument into two cases, dependent on which term is larger in the second component of the minimum.

Case 1: $N \leq T_0^{1+2\epsilon}$.

In this case we have

$$N^{2-2\beta} T_0^\epsilon + T_0^{1+\epsilon} N^{1-2\beta} \ll T_0^{1+3\epsilon} N^{1-2\beta}.$$

Hence

$$\begin{aligned} R &\ll \min(T_0^{1+\epsilon} M^{1-2\alpha}, T_0^{1+3\epsilon} N^{1-2\beta}) \\ &\ll (T_0^{1+\epsilon} M^{1-2\alpha})^{1/2} (T_0^{1+3\epsilon} N^{1-2\beta})^{1/2} \ll T_0 x_1^{1/2-\sigma+2\epsilon}. \end{aligned}$$

But for $\sigma \leq 3/4$ we have $\frac{1}{2} - \sigma \leq \frac{5}{4} - 2\sigma$. Thus

$$R \ll T_0 x_1^{5/4-2\sigma+2\epsilon}.$$

Case 2: $N > T_0^{1+2\epsilon}$.

We have by Lemmas 5.2 and 5.4

$$R \ll \min(N^{2-2\beta+\epsilon}, T_0 M^{1-2\alpha+\epsilon}),$$

$$R^* \ll M^{1-2\alpha} T_0^\epsilon (RM + R^2 + R^{5/4} T_0^{1/2})^{1/2} (R^* M + R^4 + R R^{*3/4} T_0^{1/2})^{1/2}.$$

But $RM \geq R^2, R^{5/4} T_0^{1/2}$ if $R \leq M, M^4 T_0^{-2}$.

But we have

$$R^2 \leq N^{2-2\beta} T_0^{1+2\epsilon} M^{1-2\alpha} = T_0^{1+2\epsilon} x_1^{2-2\sigma} M^{-1}$$

Thus $R \leq M T_0^\epsilon$ if $M \geq x_1^{2(1-\sigma)/3} T_0^{1/3}$ and $R \leq M^4 T_0^{-2+\epsilon}$ if $M \geq x_1^{2(1-\sigma)/9} T_0^{5/9}$. Hence, if

$$M \geq \max(x_1^{2(1-\sigma)/9} T_0^{5/9}, x_1^{2(1-\sigma)/3} T_0^{1/3})$$

then

$$RM + R^2 + R^{5/4}T_0^{1/2} \ll RMT_0^{\epsilon}$$

so

$$R^* \ll T_0^{3\epsilon}(M^{4-4\alpha}R + M^{3/2-2\alpha}R^{5/2} + M^{12/5-16\alpha/5}T_0^{2/5}R^{8/5}).$$

We now consider separately each of the three terms dominating.

Case 2A: $R^* \ll T_0^{3\epsilon}M^{4-4\alpha}R$.

$$RR^* \ll T_0^{3\epsilon}M^{4-4\alpha}R^2 \ll T_0^{5\epsilon}M^{4-4\alpha}N^{4-4\beta} \ll x_1^{4-4\sigma+5\epsilon}.$$

Since $\sigma \le 3/4$, $\mu \le 2$ we have

$$(2\sigma - 1/2)\mu \le 2.$$

Thus

$$RR^* \ll T_0^2 x_1^{9/2-6\sigma+5\epsilon}.$$

It follows that either

$$R \ll T_0 x_1^{5/4-2\sigma+2\epsilon}$$

or

$$R^* \ll T_0 x_1^{13/4-4\sigma+3\epsilon}.$$

Case 2B: $R^* \ll T_0^{3\epsilon}M^{3/2-2\alpha}R^{5/2}$.

$$\begin{aligned} R^* &\ll T_0^{3\epsilon}M^{3/2-2\alpha}R^{5/2} \\ &\ll T_0^{4\epsilon}M^{3/2-2\alpha}N^{2-2\beta}\left(N^{2-2\beta}T_0^{1+2\epsilon}M^{1-2\alpha}\right)^{3/4} \\ &\ll x_1^{7(1-\sigma)/2+6\epsilon}M^{-5/4}T_0^{3/4}. \end{aligned}$$

But then for $M \ge x_1^{(2\sigma+1)/5}T_0^{-1/5}$ we have

$$R^* \ll T_0 x_1^{13/4-4\sigma+6\epsilon}.$$

Case 2C: $R^* \ll M^{12/5-16\alpha/5}R^{8/5}T_0^{2/5+3\epsilon}$.

$$R^* \ll M^{12/5-16/5\alpha}N^{16/5-16/5\alpha}T_0^{2/5+5\epsilon} \ll x_1^{16(1-\sigma)/5+5\epsilon}T_0^{2/5}M^{-4/5}$$

But then for $M > x_1^{\sigma-1/16}T_0^{-3/4}$ we have

$$R^* \ll T_0 x_1^{13/4-4\sigma+5\epsilon}.$$

Therefore the Lemma holds, provided that we can always combine polynomials to ensure that

$$M > x_1^{\sigma-1/16}T_0^{-3/4}, x_1^{(2\sigma+1)/5}T_0^{-1/5}, x_1^{2(1-\sigma)/9}T_0^{5/9}, x_1^{2(1-\sigma)/3}T_0^{1/3}.$$

We claim that we can always combine polynomials to ensure that the smaller polynomial M satisfies $M \ge \min(x_1^{2/5}, x_1/T_0^{1+2\epsilon})$. It suffices to find a product P of polynomials with length in the interval $[\min(x_1/T_0^{1+2\epsilon}, x_1^{2/5}), \max(T_0^{1+2\epsilon}, x_1^{3/5})]$ since then either P or the complementary product will have suitable length. To obtain P we combine polynomials S_i which are not the exceptional polynomial in decreasing order of length until we find the first product, $S^{(1)}S^{(2)}\ldots S^{(r)}$ say, with length $\ge \min(x_1/T_0^{1+2\epsilon}, x_1^{2/5})$. Since the exceptional polynomial has length $\le x_1^{3/5}$ such a product exists. We let $S^{(i)}$ have length L_i. Therefore $L_1\ldots L_r > \min(x_1/T_0^{1+2\epsilon}, x_1^{2/5})$ and so we have found a suitable product unless $L_1\ldots L_r > \max(T_0^{1+2\epsilon}, x_1^{3/5})$. Since $L_i \le T_0^{1/2+\epsilon}$ for all i this means we must have $r \ge 3$. By construction we must also have that $L_1\ldots L_{r-1} < \min(x_1/T_0^{1+2\epsilon}, x_1^{2/5})$ and so

$L_r \geq (L_1 \dots L_r)/(L_1 \dots L_{r-1}) \geq x_1^{1/5}$. Since by construction $L_i \geq L_r \geq x_1^{1/5}$ for all $i < r$ we have that $L_1 \dots L_{r-1} \geq x_1^{(r-1)/5} \geq x_1^{2/5}$. But this is a contradiction with $L_1 \dots L_{r-1} < \min(x_1/T_0^{1+2\epsilon}, x_1^{2/5})$, and so we must have that $L_1 \dots L_r \in [\min(x_1/T_0^{1+2\epsilon}, x_1^{2/5}), \max(T_0^{1+2\epsilon}, x_1^{3/5})]$.

Since in the case we are considering $N > T_0^{1+2\epsilon}$, we have $M < x_1/T_0^{1+2\epsilon}$. We also have that $M \geq \min(x_1/T_0^{1+2\epsilon}, x_1^{2/5})$ by the above construction. Therefore we must have $\mu > 5/3$. For $\mu > 5/3$ and $\sigma \geq 0.7$, we have

$$M \geq x_1^{2/5} \geq x_1^{11/16}T_0^{-3/4}, x_1^{1/2}T_0^{-1/5}, x_1^{2(1-\sigma)/9}T_0^{5/9}, x_1^{2(1-\sigma)/3}T_0^{1/3}$$

and so the Lemma holds. ◻

5.4. **Part 3:** $3/4 \leq \sigma \leq 1$, μ **small.**

We now consider the range $3/4 \leq \sigma \leq 1$, $\mu \leq 4/(4\sigma - 1) + \epsilon$.

Lemma 5.9. *Let $3/4 \leq \sigma$ and $4/3 \leq \mu \leq \frac{4}{4\sigma-1} + \epsilon$. Then we have*

$$R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$$

or

$$R \ll T_0 x_1^{5/4-2\sigma+2\epsilon}$$

or

$$R^* \ll T_0 x_1^{13/4-4\sigma+8\epsilon}.$$

Proof. We assume that $R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$ does not hold. Therefore the results of Lemma 5.6 apply.

The result follows from published estimates if $\sigma \geq 13/16$ or if $\mu \leq 8/5$.

By (139) (choosing $\delta = \epsilon$) we have

$$R^* \ll T_0^{12(1-\sigma)/(4\sigma-1)+\epsilon} \ll T_0 T_0^{(13-16\sigma)/(4\sigma-1)+\epsilon}.$$

This gives $R^* \ll T_0 x_1^{13/4-4\sigma+\epsilon}$ for $\sigma \geq 13/16$ since $\mu \leq 4/(4\sigma - 1) + \epsilon$. Thus without loss of generality we assume $\sigma \leq 13/16$.

By (135) (choosing $\delta = \epsilon$) we have

$$R \ll T_0^{3(1-\sigma)/(3\sigma-1)+\epsilon}.$$

This gives $R \ll T_0 x_1^{5/4-2\sigma+\epsilon}$ provided that

$$\mu \leq \frac{6\sigma - 4}{(3\sigma - 1)(2\sigma - 5/4)}.$$

For $3/4 \leq \sigma \leq 13/16$ this covers the range $\mu \leq 8/5$. Therefore without loss of generality we assume $\mu \geq 8/5$.

We now consider the remaining range $3/4 \leq \sigma \leq 13/16$ and $\mu \geq 8/5$ in the same manner as our argument in Part 2.

We combine the polynomials into two polynomials M, N as in Lemma 5.8. Therefore we can choose M such that $\min(x_1 T_0^{-1-2\epsilon}, x_1^{2/5}) \leq M \leq N$.

We use Lemma 5.3 (choosing $\delta = \epsilon$) to give

$$R \ll \min(T_0^{\epsilon} M^{2-2\alpha} + T_0^{1+\epsilon} M^{4-6\alpha}, T_0^{\epsilon} N^{2-2\beta} + T_0^{1+\epsilon} N^{4-6\beta}).$$

We split our argument into four cases, dependent on which terms dominate in this estimate.

Case 1: $N^{4\beta-2} \le T_0$, $M^{4\alpha-2} \le T_0$.

$$\begin{aligned} R &\ll \min(T_0^{1+\epsilon}M^{4-6\alpha}, T_0^{1+\epsilon}N^{4-6\beta}) \\ &\ll (T_0^{1+\epsilon}M^{4-6\alpha})^{1/2}(T_0^{1+\epsilon}N^{4-6\beta})^{1/2} \\ &\ll T_0 x_1^{2-3\sigma+\epsilon}. \end{aligned}$$

But for $\sigma \ge 3/4$, $2-3\sigma \le \frac{5}{4} - 2\sigma$ and so

$$R \ll T_0 x_1^{5/4-2\sigma+\epsilon}.$$

Case 2: $N^{4\beta-2} \le T_0$, $M^{4\alpha-2} > T_0$.

$$M^{1-2\alpha} = x_1^{1-2\sigma}(N^{4\beta-2})^{1/2} \le T_0^{1/2}x_1^{1-2\sigma}.$$

Hence

$$R \ll T_0^{\epsilon}M^{2-2\alpha} \le T_0^{1/2+\epsilon}x_1^{1-2\sigma}M \le T_0 x_1^{5/4-2\sigma+\epsilon}$$

since $M \le x_1^{1/2} \le T_0$.

Case 3: $N^{4\beta-2} > T_0$, $M^{4\alpha-2} > T_0$.

$$R \ll (M^{2-2\alpha+\epsilon}N^{2-2\beta+\epsilon})^{1/2} = x_1^{1-\sigma+\epsilon}.$$

But, for $\mu \le 4/(4\sigma-1)+\epsilon$, we have:

$$R \ll x_1^{1-\sigma+\epsilon} \le T_0 x_1^{5/4-2\sigma+2\epsilon}.$$

Case 4: $N^{4\beta-2} > T_0$, $M^{4\alpha-2} \le T_0$.

By Lemmas 5.3 and 5.4 we have

$$R \ll \min(T_0^{\epsilon}N^{2-2\beta}, T_0^{1+\epsilon}M^{4-6\alpha}),$$

$$R^* \ll M^{1-2\alpha}T_0^{\epsilon}(RM + R^2 + R^{5/4}T_0^{1/2})^{1/2}(R^*M + R^4 + RR^{*3/4}T_0^{1/2})^{1/2}.$$

But $RM \ge R^2, R^{5/4}T_0^{1/2}$ if $R \le M, M^4T_0^{-2}$.

We have

$$R^4 \ll N^{6-6\beta}T_0^{1+4\epsilon}M^{4-6\alpha} = T_0^{1+4\epsilon}x_1^{6-6\sigma}M^{-2}.$$

Thus $R^2 \ll RMT_0^{\epsilon}$ if $M \ge x_1^{1-\sigma}T_0^{1/6}$ and $R^{5/4}T_0^{1/2} \ll RMT_0^{3\epsilon}$ if $M \ge x_1^{(1-\sigma)/3}T_0^{1/2-2\epsilon}$. Hence, if

$$M \ge \max(x_1^{(1-\sigma)/3}T_0^{1/2-2\epsilon}, x_1^{1-\sigma}T_0^{1/6})$$

then

$$R^2 + R^{5/4}T_0^{1/2} + RM \ll RMT_0^{3\epsilon}.$$

In this case

$$R^* \ll T_0^{5\epsilon}(M^{4-4\alpha}R + M^{3/2-2\alpha}R^{5/2} + M^{12/5-16\alpha/5}T_0^{2/5}R^{8/5}).$$

We now consider separately each term dominating the RHS.

Case 4A: $R^* \ll T_0^{5\epsilon}M^{4-4\alpha}R$.

$$RR^* \ll T_0^{5\epsilon}M^{4-4\alpha}R^2 \ll T_0^{7\epsilon}M^{4-4\alpha}N^{4-4\beta} \ll x_1^{4-4\sigma+7\epsilon}.$$

Since $\mu \le 4/(4\sigma-1)+\epsilon$ we have

$$RR^* \ll x_1^{4-4\sigma+7\epsilon} \ll T_0^2 x_1^{9/2-6\sigma+9\epsilon}.$$

It follows that either

$$R \ll T_0 x_1^{5/4-2\sigma+2\epsilon}$$

or

$$R^* \ll T_0 x_1^{13/4-4\sigma+7\epsilon}.$$

Case 4B: $R^* \ll T_0^{5\epsilon} M^{3/2-2\alpha} R^{5/2}$.

$$\begin{aligned} R^* &\ll T_0^{5\epsilon} M^{3/2-2\alpha} R^{5/2} \\ &\ll T_0^{6\epsilon} M^{3/2-2\alpha} N^{2-2\beta} \left(T_0^{1+4\epsilon} N^{6-6\beta} M^{4-6\alpha}\right)^{3/8} \\ &\ll x_1^{17(1-\sigma)/4+8\epsilon} M^{-5/4} T_0^{3/8}. \end{aligned}$$

Hence for $M > T_0^{-1/2} x_1^{(4-\sigma)/5}$ we have

$$R^* \ll T_0 x_1^{13/4-4\sigma+8\epsilon}.$$

Case 4C: $R^* \ll M^{12/5-16\alpha/5} R^{8/5} T_0^{2/5+5\epsilon}$.

$$\begin{aligned} R^* &\ll M^{12/5-16\alpha/5} R^{8/5} T_0^{2/5+5\epsilon} \\ &\ll M^{12/5-16/5\alpha} N^{16/5-16/5\beta} T_0^{2/5+7\epsilon} \\ &\ll x_1^{16(1-\sigma)/5+7\epsilon} T_0^{2/5} M^{-4/5}. \end{aligned}$$

But for $M > T_0^{-3/4} x_1^{\sigma-1/16}$ we have

$$\ll T_0 x_1^{13/4-4\sigma+7\epsilon}.$$

Therefore the Lemma holds, provided that we can always combine the polynomials to ensure that

$$M > x_1^{1-\sigma} T_0^{1/6}, x_1^{(1-\sigma)/3} T_0^{1/2-2\epsilon}, T_0^{-1/2} x_1^{(4-\sigma)/5}, T_0^{-3/4} x_1^{\sigma-1/16}.$$

For $3/4 \le \sigma \le 13/16$, $8/5 \le \mu \le 4/(4\sigma-1)$ we have

$$x_1^{2/5} \ge x_1^{1-\sigma} T_0^{1/6}, x_1^{(1-\sigma)/3} T_0^{1/2}, T_0^{-1/2} x_1^{(4-\sigma)/5}, T_0^{-3/4} x_1^{\sigma-1/16}$$

and

$$\frac{x_1}{T_0^{1+2\epsilon}} \ge x_1^{1-\sigma} T_0^{1/6}, T_0^{-1/2} x_1^{(4-\sigma)/5}, T_0^{-3/4} x_1^{\sigma-1/16}$$

If

$$\mu \ge \frac{9}{4+2\sigma}$$

then we have

$$\frac{x_1}{T_0^{1+2\epsilon}} \ge x_1^{(1-\sigma)/3} T_0^{1/2-2\epsilon}.$$

and so

$$M > \min(x_1 T_0^{-1}, x_1^{2/5}) > x_1^{1-\sigma} T_0^{1/6}, x_1^{(1-\sigma)/3} T_0^{1/2-2\epsilon}, T_0^{-1/2} x_1^{(4-\sigma)/5}, T_0^{-3/4} x_1^{\sigma-1/16}$$

as required.

We therefore consider $\mu \le 9/(4+2\sigma)$. Since we are considering $N^{4\beta-2} > T_0$, if $N \le T_0^{3/2+2\epsilon} x_1^{5/4-2\sigma}$ then

$$R \ll T_0^{\epsilon} N^{2-2\beta} \ll \frac{N}{T_0^{1/2-\epsilon}} \ll T_0 x_1^{5/4-2\sigma+3\epsilon}.$$

Therefore we only need to consider $N \ge T_0^{3/2+2\epsilon} x_1^{5/4-2\sigma}$, and so (since $NM = x_1$)

$$\frac{x_1}{T_0^{1+2\epsilon}} \le M \le \frac{x_1^{2\sigma-1/4}}{T_0^{3/2+2\epsilon}}.$$

This means we must have
$$\mu \geq \frac{2}{8\sigma - 5}.$$
Since we also have
$$\mu \leq \frac{9}{4 + 2\sigma}$$
we must have $\sigma \geq 53/68$. But in this range we can use (135) again. This gives
$$R \ll T_0^{3(1-\sigma)/(3\sigma-1)+\epsilon} \ll T_0 x_1^{5/4-2\sigma+\epsilon}$$
provided we have
$$\mu \leq \frac{8(3\sigma - 2)}{(8\sigma - 5)(3\sigma - 1)}.$$
This, combined with $\mu \leq 9/(4 + 2\sigma)$ means that we must have $\sigma \leq (271 - \sqrt{193})/336 < 53/68$. Therefore we have covered all possible values of $\mu \geq 8/5$ and $3/4 \leq \sigma \leq 13/16$. Thus the Lemma holds. □

5.5. **Part 4:** $3/4 \leq \sigma \leq 1-10^{-22}$**,** μ **large.** We now consider the range of $\mu \geq 4/(4\sigma-1)+\epsilon$, $3/4 \leq \sigma \leq 1 - 10^{-22}$. We require separate treatment for σ very close to 1.

Lemma 5.10. *Let* $3/4 \leq \sigma \leq 1 - 10^{-22}$, $\mu \geq 4/(4\sigma - 1) + \epsilon$. *Then either*
$$R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$$
or
$$R^* \ll T_0 x_1^{13/4-4\sigma+8\epsilon}.$$

Proof. We assume that $R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$ does not hold. Therefore the results of Lemmas 5.5 and 5.6 apply.

The result follows from published estimates unless $13/16 \leq \sigma \leq 25/28$ and $\mu \leq 3/(10\sigma - 7) + \epsilon$.

By (138) (choosing $\delta = \epsilon$) we have
$$R^* \ll T_0^{(12-12\sigma)/(4\sigma-1)+\epsilon} = T_0 T_0^{(13-16\sigma)/(4\sigma-1)+\epsilon}.$$
Since we have $\mu \geq 4/(4\sigma - 1) + \epsilon$, if $\sigma \leq 13/16$ this gives $R^* \ll T_0 x_1^{13/4-4\sigma+\epsilon}$. Therefore without loss of generality we may assume $13/16 \leq \sigma$.

By (139), if $\sigma \geq 25/28$ then (choosing $\delta = 10^{-23}\epsilon$) we have that
$$R \ll T_0^{(4-4\sigma)/(4\sigma-1)+10^{-23}\epsilon}.$$
Since $\mu \geq 4/(4\sigma - 1) + \epsilon$ and $1 - \sigma \geq 10^{-22}$, we have for $\sigma \geq 25/28$ that
$$R \ll T_0^{(4/(4\sigma-1)+\epsilon)(1-\sigma)-\epsilon(1-\sigma)+10^{-23}\epsilon} \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}.$$
Therefore without loss of generality we may assume $\sigma \leq 25/28$.

By (137) (choosing $\delta = \epsilon/28$) we have
$$R \ll T_0^{(3-3\sigma)/(10-7\sigma)+\epsilon/28}.$$
Therefore if $\mu \geq 3/(10\sigma - 7) + \epsilon$ and $13/16 \leq \sigma \leq 25/28$ we have
$$R \ll T_0^{(3/(10\sigma-7)+\epsilon)(1-\sigma)-\epsilon(1-\sigma)+\epsilon/28} \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}.$$
Therefore without loss of generality we may assume $\mu \leq 3/(10\sigma - 7) + \epsilon$.

We now consider the remaining case of $13/16 \le \sigma \le 25/28$ and $\mu \le 3/(10\sigma - 7) + \epsilon$.

We repeatedly combine any pair of polynomials of length $\le T_0^{10^{-24}\epsilon}$ so at most one polynomial has length $\le T_0^{10^{-24}\epsilon}$.

We then pick a polynomial S_{j_1} of length $N_{j_1} \ge T_0^{10^{-24}\epsilon}$ with σ_{j_1} maximal.

If $\sigma_{j_1} \le \sigma$ then, since σ is an average of the σ_i, the polynomial S_{j_2} with length $N_{j_2} \le T_0^{10^{-24}\epsilon}$ must exist and have $\sigma_{j_2} \ge \sigma$. In this case we combine the polynomial S_{j_1} and S_{j_2} to produce a polynomial S of length N and size $\sigma' \ge \sigma$.

If $\sigma_{j_1} \ge \sigma$ then we take $S = S_{j_1}$, (and so $N = N_{j_1}$, $\sigma' = \sigma_{j_1}$).

We consider separately the cases when N is small and N is large.

Case 1:$N \le T_0^{1/(4\sigma-1)}$.

We follow the analysis of Heath-Brown in [13][Pages 228-230].

If $N \le T_0^{1/(8\sigma-2)}$ we raise it to a suitable exponent so that the new polynomial has length M with $T_0^{1/(8\sigma-2)} \le M \le T_0^{1/(4\sigma-1)}$. We note that since $N \ge T_0^{10^{-22}\epsilon}$ this exponent is $O(1)$ and so all the coefficients are still $O_\delta(T_0^\delta)$ for every $\delta > 0$.

We raise M to different exponents to use in the R and R^* estimates. We let $M_1 = M^{k_1}$ which we will use for bound R and we let $M_2 = M^{k_2}$ which we will use to bound R^*. We choose k_2 such that $M^{k_2} \le T_0^{2/(4\sigma-1)} < M^{1+k_2}$, which means $k_2 = 2$ or 3 and $T_0^{4/(12\sigma-3)} \le M_2 \le T_0^{2/(4\sigma-1)}$. We pick $k_1 = k_2$ when $T_0^{1/(3\sigma-1)} \le M_2$ and $k_1 = 1 + k_2$ for $M_2 \le T_0^{1/(3\sigma-1)}$. Using Lemma 5.3 and recalling that $\sigma' \ge \sigma$ this gives for any $\delta > 0$

$$R \ll_\delta \begin{cases} T_0^\delta M_1^{2-2\sigma}, & T_0^{1/(4\sigma-2)} \le M_1 \\ T_0^{1+\delta} M_1^{4-6\sigma}, & M_1 \le T_0^{1/(4\sigma-2)}. \end{cases}$$

If $M_2 \le T_0^{1/(3\sigma-1)}$ then $M_1 = M_2^{4/3}$ or $M_2^{3/2}$ (depending on whether $k_1 = 2$ or 3). In this case, for either value of k_1, we get $R \ll_\delta T_0^\delta(TM_2^{16/3-8\sigma} + M_2^{3-3\sigma}) \ll T_0^\delta M_2^{3-3\sigma}$. Using this bound is sufficient for our purposes.

Thus, using the above and Lemma 5.3 we get the following bound for any $\delta > 0$

$$R \ll_\delta \begin{cases} T_0^\delta M_2^{2-2\sigma}, & T_0^{1/(4\sigma-2)} \le M_2 \le T_0^{2/(4\sigma-1)} \\ T_0^{1+\delta} M_2^{4-6\sigma}, & T_0^{1/(3\sigma-1)} \le M_2 \le T_0^{1/(4\sigma-2)} \\ T_0^\delta M_2^{3-3\sigma}, & T_0^{4/(12\sigma-3)} \le M_2 \le T_0^{1/(3\sigma-1)}. \end{cases}$$

We now consider each range of M_2 separately.

Case 1A: $T_0^{1/(4\sigma-2)} \le M_2 \le T_0^{2/(4\sigma-1)}$.

For $T_0^{1/(4\sigma-2)} \le M_2 \le T_0^{2/(4\sigma-1)}$ we have for $\delta = \epsilon/28$

$$R \ll M^{2-2\sigma} T_0^{\epsilon/28} \ll T_0^{(4/(4\sigma-1)+\epsilon)(1-\sigma)+\epsilon/28-\epsilon(1-\sigma)} \ll x_1^{1-\sigma} (\log x_1)^{-4k-2}$$

since $1 - \sigma \ge 3/28$ and $\mu \ge 4/(4\sigma - 1) + \epsilon$.

Case 1B: $T_0^{1/(3\sigma-1)} \le M_2 \le T_0^{1/(4\sigma-2)}$.

For $T_0^{1/(3\sigma-1)} \leq M_2 \leq T_0^{1/(4\sigma-2)}$ we have $R \ll T_0^{1+\epsilon} M_2^{4-6\sigma}$. This means that $R^{5/4}T_0^{1/2} \ll RM_2$ and $R^2 \ll RM_2$, so Lemma 5.4 simplifies to

$$R^* \ll T_0^{2\epsilon}(RM_2^{4-4\sigma} + R^{5/2}M_2^{(3-4\sigma)/2} + R^{8/5}T_0^{2/5}M_2^{(12-16\sigma)/5}).$$

Using $R \ll T_0^{1+\epsilon}M_2^{4-6\sigma}$ this gives

$$R^* \ll T_0^{5\epsilon}(T_0 M_2^{8-10\sigma} + T_0^{5/2}M_2^{(23-34\sigma)/2} + T_0^2 M_2^{(44-64\sigma)/5}).$$

Since we are considering $\sigma \geq 13/16$, all the exponents of M_2 are negative. Thus since $M_2 \geq T_0^{1/(3\sigma-1)}$ we have

$$R^* \ll T_0^{5\epsilon}(T_0^{(7-7\sigma)/(3\sigma-1)} + T_0^{(18-19\sigma)/(6\sigma-2)} + T_0^{(34-34\sigma)/(15\sigma-5)}).$$

But for $\sigma \geq 13/16$ we have

$$\max\left(\frac{7-7\sigma}{3\sigma-1}, \frac{18-19\sigma}{6\sigma-2}, \frac{34-34\sigma}{15\sigma-5}\right) \leq 1 + \left(\frac{3}{10\sigma-7}\right)\left(\frac{13-16\sigma}{4}\right).$$

This means that we have $R \ll T_0 x_1^{13/4-4\sigma+6\epsilon}$ since $\mu \leq 3/(10\sigma-7)+\epsilon$.

Case 1C: $T_0^{4/(12\sigma-3)} \leq M_2 \leq T_0^{1/(3\sigma-1)}$.

For $T_0^{4/(12\sigma-3)} \leq M_2 \leq T_0^{1/(3\sigma-1)}$ we have $R \ll M_2^{3-3\sigma}T_0^{\epsilon}$. Therefore $R \ll M_2$ and Lemma 5.4 simplifies to

$$\begin{aligned}R^* \ll M_2^{1-2\sigma}T_0^{\epsilon}(&R^{1/2}R^{*1/2}M_2 + R^{5/2}M_2^{1/2} + RR^{*3/8}M_2^{1/2}T_0^{1/4}\\ &+ R^{5/8}R^{*1/2}M_2^{1/2}T_0^{1/4} + R^{21/8}T_0^{1/4} + R^{9/8}R^{*3/8}T_0^{1/2}).\end{aligned}$$

We note that using the trivial bound $R^* \ll (\log T)R^3$ we have $RR^{*3/8}M_2^{1/2}T_0^{1/4+\epsilon} \gg R^{5/8}R^{*1/2}M_2^{1/2}T_0^{1/4}$. We can therefore drop the fourth term at the cost of a factor $\ll T_0^{\epsilon}$. This yields

$$\begin{aligned}R^* \ll T_0^{4\epsilon}(&RM_2^{4-4\sigma} + R^{5/2}M_2^{(3-4\sigma)/2} + R^{8/5}M_2^{(12-16\sigma)/5}T_0^{2/5}\\ &+ R^{21/8}T_0^{1/4}M_2^{1-2\sigma} + R^{9/5}M_2^{(8-16\sigma)/5}T_0^{4/5}).\end{aligned}$$

Substituting in $R \ll M_2^{3-3\sigma+\epsilon}$ we get

$$\begin{aligned}R^* \ll T_0^{7\epsilon}(&M_2^{7-7\sigma} + M_2^{(18-19\sigma)/2} + M_2^{(36-40\sigma)/5}T_0^{2/5}\\ &+ M_2^{(71-79\sigma)/8}T_0^{1/4} + T_0^{4/5}M_2^{(35-43\sigma)/5}).\end{aligned}$$

For $13/16 \leq \sigma \leq 25/28$ the first four terms always have positive exponents of M_2. Thus, using $T_0^{4/(12\sigma-3)} \leq M_2 \leq T_0^{1/(3\sigma-1)}$ we get

$$\begin{aligned}R^* \ll T_0^{7\epsilon}(&T_0^{(7-7\sigma)/(3\sigma-1)} + T_0^{(18-19\sigma)/(6\sigma-2)} + T_0^{(34-34\sigma)/(15\sigma-5)}\\ &+ T_0^{(69-73\sigma)/(24\sigma-8)} + T_0^{(31-31\sigma)/(15\sigma-5)} + T_0^{(128-124\sigma)/(60\sigma-15)}).\end{aligned}$$

But for $13/16 \leq \sigma \leq 25/28$ we have

$$\begin{aligned}\max\Bigg(&\frac{7-7\sigma}{3\sigma-1}, \frac{18-19\sigma}{6\sigma-2}, \frac{34-34\sigma}{15\sigma-5},\\ &\frac{69-73\sigma}{24\sigma-8}, \frac{31-31\sigma}{15\sigma-5}, \frac{128-124\sigma}{60\sigma-15}\Bigg) \leq 1 + \left(\frac{13-16\sigma}{4}\right)\left(\frac{3}{10\sigma-7}\right).\end{aligned}$$

Thus $R^* \ll T_0 x_1^{13/4-4\sigma+8\epsilon}$ since $\mu \leq \frac{3}{10\sigma-7}+\epsilon$.

Putting these estimates together covers all possible values of $\mu \geq 4/(4\sigma-1)+\epsilon$ and $3/4 \leq \sigma \leq 1-10^{-22}$.

Case 2: $N \geq T_0^{1/(4\sigma-1)}$.

We will choose $k \geq 7$ so that (23) implies that for $i \leq k$ we have
$$N_i \leq (3x)^{1/7} \leq T_0^{1/3} \leq T_0^{1/(4\sigma-1)}.$$
Therefore the polynomial selected with this 'long' length must either be one with all coefficients 1 or $\log(n)$ and length $N = N_{j_1}$, or it must be a the combination of such a polynomial with another polynomial of length $N_{j_2} \leq T_0^{10^{-24}\epsilon}$ (so $N = N_{j_1}N_{j_2}$).

By Lemma 5.5 either
$$R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$$
or
$$R \ll_\delta T_0^{2+\delta} N_{j_1}^{6-12\sigma_{j_1}}$$
for any $\delta > 0$.

Without loss of generality we assume $R \ll_\delta T_0^{2+\delta} N_{j_1}^{6-12\sigma_{j_1}}$.

If $N = N_{j_1}N_{j_2}$ then since $N_{j_2} \ll T_0^{10^{-24}\epsilon}$ we have
$$\begin{aligned} R &\ll R(T_0^{6\times 10^{-24}\epsilon} N_k^{6-12\sigma_k}) \\ &\ll T_0^{2+10^{-23}\epsilon} N^{6-12\sigma'}. \end{aligned}$$
The same result clearly holds if $N = N_{j_1}$.

Thus, since $N \geq T_0^{1/(4\sigma-1)}$, we have
$$\begin{aligned} R &\ll T_0^{2-6(2\sigma-1)/(4\sigma-1)+10^{-23}\epsilon} \\ &\ll T_0^{(1-\sigma)(4/(4\sigma-1)+\epsilon)-\epsilon(1-\sigma)+10^{-23}\epsilon} \\ &\ll x_1^{(1-\sigma)}(\log x_1)^{-4k-2} \end{aligned}$$
since
$$\mu > \frac{4}{4\sigma-1} + \epsilon, \qquad 1-\sigma \geq 10^{-22}.$$
□

5.6. **Part 5:** $1-10^{-22} \leq \sigma \leq 1$**.** We now consider the final range, when $1-10^{-22} \leq \sigma \leq 1$. We split the argument into two cases - when σ is exceptionally close to one, and so we can use Vinogradov's bound, and the remaining case.

Lemma 5.11. *Let* $1 - 10^{-22} \leq \sigma \leq 1$, $\mu \geq 4/(4\sigma-1) + \epsilon$. *Then we have*
$$R \ll x_1^{1-\sigma}(\log x_1)^{-4k-2}$$

Proof. We recall that
$$\eta = \eta(x_1) = C_0(\log x_1)^{-2/3}(\log\log x_1)^{-1/3}$$
for some constant $C_0 > 0$ from (132).

Case 1: $\sigma \leq 1 - \eta$

We consider separately the case when all polynomials are small.

Case 1A: $N_i \leq x_1^{1/k}\ \forall\ i$

We repeatedly combine pairs of polynomials of length $\leq x_1^{1/2k}$. Thus without loss of generality we assume all polynomials have length $\in [x_1^{1/2k}, x_1^{1/k}]$, except for possibly one polynomial with length $\leq x_1^{1/2k}$. We combine this polynomial with one of the remaining ones, so all polynomials have length $\in [x_1^{1/2k}, x_1^{3/2k}]$.

Since we have combined at most $2k$ polynomials, all remaining polynomials must have coefficients which are $\ll (\log x_1) d_{2k}(n)$.

We pick a polynomial which has $\sigma_i \geq \sigma$. We raise this polynomial to an exponent so that it has length Y with $T_0^{9/16} \leq Y \leq T_0^{5/8}$. This is possible if the polynomial's original length was $\leq T_0^{1/16}$. This is the case provided $x_1^{3/2k} \leq T_0^{1/16}$. Thus if we now choose

$$k = 60 \tag{143}$$

then this is satisfied, since $\mu \leq 19/9$. We note that we raise the polynomial to an exponent ≤ 75, and so all the coefficients of this new polynomial of length Y are $\ll (\log x_1)^{75} d_{9000}(n)$.

Using Lemma 5.3 we have

$$\begin{aligned} R &\ll (\log x_1)^2 (Y^{2-2\sigma_i} + T_0 Y^{4-6\sigma_i}) \left(1 + \frac{\sum_Y^{2Y} (\log x_1)^{150} d_{9000}(n)^2}{Y}\right)^3 \\ &\ll (\log x_1)^{60000} (Y^{2-2\sigma} + T_0 Y^{4-6\sigma}) \end{aligned}$$

But for $\sigma \geq 17/18$ we have $T_0^{9(4\sigma-2)/16} \geq T_0$ and so $Y^{4\sigma-2} \geq T_0$. This means that $Y^{2-2\sigma} \geq T_0 Y^{4-6\sigma}$. Hence

$$\begin{aligned} R &\ll Y^{2-2\sigma} (\log x_1)^{60000} \\ &\ll T_0^{5(1-\sigma)/4} (\log x_1)^{60000} \\ &\ll x_1^{15(1-\sigma)/16} (\log x_1)^{60000} \\ &\ll x_1^{1-\sigma} \exp(-C_0 (\log x_1)^{1/3} (\log\log x_1)^{-2/3}/16 + 60000 \log\log x_1) \\ &\ll x_1^{1-\sigma} (\log x_1)^{-4k-2} \end{aligned}$$

Since $\sigma \leq 1 - \eta$, $\mu \geq 4/3$.

Case 1B: $\exists\, N_j > x_1^{1/k}$

By (23) this polynomial must have all coefficients 1 or $\log(n)$. We first consider the case when all coefficients are 1.

By Van-der-Corput's method of exponential sums (see the proof of [30][Theorem 5.14], for example) we have for any $l \in \mathbb{Z}$ and $\alpha = 1 - l/(2^l - 2)$

$$\sum_N^{2N} n^{-\alpha - it} \ll T_0^{1/(2^l-2)+\epsilon}$$

Thus

$$\begin{aligned} |S_j| = \left| \sum_{N_j}^{2N_j} n^{-c-it} \right| &\ll N_j^{-c+\alpha} T_0^{\epsilon} \sup_{N \in [N_j/2, N_j]} \left| \sum_N^{2N} n^{-\alpha - it} \right| \\ &\ll N_j^{-l/(2^l-2)} T_0^{1/(2^l-2)+2\epsilon}. \end{aligned}$$

The same result folds for a polynomial with all coefficients $(\log n)$ by partial summation.

We have $N_j > x_1^{1/k} > T_0^{1/k}$ so

$$|S_j| \ll N_j^{-(l-k)/(2^l-2)+2k\epsilon}$$

Thus, choosing $l = k+1$ we have

$$|S_j| \ll N_j^{-2^{-k-1}+2k\epsilon}$$

and so (for $\epsilon \leq 2^{-k-3}/k$) we must have

$$\sigma_j \leq 1 - 2^{-k-1} + 2k\epsilon \leq 1 - 2^{-k-2}$$

since $|S_i| \ll N_i^{-1+\sigma_i}$. But

$$\begin{aligned} x_1^{\sigma} &= \prod_i N_i^{\sigma_i} \\ &\ll N_j^{\sigma_j} \prod_{i \neq j} N_i \\ &= x_1 N_j^{\sigma_j - 1} \\ &\ll x_1 (x_1^{1/k})^{-2^{-k-2}} \\ &\ll x_1^{1-2^{-k-2}/k}. \end{aligned}$$

Hence

$$\sigma \leq 1 - \frac{1}{k2^{k+2}}$$

or $R = 0$.

By (143) we have $k = 60$, so this means that

$$\sigma \leq 1 - 10^{-22} \qquad \text{or} \qquad R = 0$$

and so we are done.

Case 2: $\sigma \geq 1 - \eta$

By the same argument as in Case 1 of Lemma 5.7, we get the result (133)

$$R = 0 \qquad \text{or} \qquad \sigma_{j'} \leq 1 - 3\eta/2$$

for any polynomial with $N_{j'} \gg x_1^{1/40k}$.

Thus either $R = 0 \ll T_0 x_1^{2\sigma - 5/4 + \epsilon}$ or (since there are only $2k$ polynomials S_i)

$$x_1^{\sigma} = \prod_{i=1}^{2k} N_i^{\sigma_i} \leq \left(\prod_{\substack{1 \leq i \leq 2k \\ N_i \leq x_1^{1/40k}}} N_i \right) \left(\prod_{\substack{1 \leq i \leq 2k \\ N_i \geq x_1^{1/40k}}} N_i^{1-3\eta/2} \right) \leq x_1^{1-\eta}$$

and so we must have $\sigma < 1 - \eta$. □

This covers all the different cases, and so the main result holds.

6. Acknowledgment

I would like to thank my supervisor, Prof. Heath-Brown, for suggesting this problem, for providing a large number of helpful comments as well as encouragement. I would also like to thank Christian Elscholtz and Yoichi Motohashi of making me aware of the existing result of Peck in this subject.

7. Comments and Further Work

Using the above argument we obtain the best possible result in some sense. Without improving the existing estimates for large values of Dirichlet Polynomials, it appears an exponent of $5/4+\epsilon$ is the smallest obtainable using the method presented.

The critical case in the argument appears to be when $\sigma_i = 3/4\ \forall i$. If there are 4 polynomials all of equal length (i.e. length $x_1^{1/4}$) then throughout the range $x_1^{11/20} \le \tau \le x_1^{5/8}$ Proposition 3.2 fails to hold for any exponent $\le 5/4$ using the estimates for the frequency of large values of Dirichlet Polynomials when $\sigma = 3/4$. In this region we use the strongest known such bounds, and so an improvement to the result would require a stronger large values estimate when $\sigma = 3/4$. Improving the estimates at $\sigma = 3/4$ appears to be difficult. Several improvements have been made to Montgomery's and Huxleys estimates given in Lemmas 5.2 and 5.3 for other ranges of σ, but $\sigma = 3/4$ appears to be the hardest to improve. The bounds given are also tight in the region $\sigma \ge 25/28$, but it appears for σ large there is more flexibility to improve the large value estimates. For σ large there are various stronger estimates for R which have not been employed here.

The bound obtained is tight in τ only for $x^{11/20} \le \tau \le x^{5/8}$ or $\tau = x^{1/2}$, which is far from the full range. Therefore the above argument implies a slightly stronger result, where we have

$$\sum_{p_n \le x} f(d_n) \ll x^{5/4+\epsilon}$$

for some function $f(t) \ge t^2$ and $f(t) \ge t^{2+\epsilon_1}$ for the range when $\tau \le x^{1/2-\epsilon'}$, or $x^{1/2+\epsilon'} \le \tau \le x^{11/20-\epsilon'}$, or $\tau \ge x^{5/8+\epsilon'}$ for some $\epsilon' > 0$.

It might be possible to improve the result by combining the method with sieve ideas. This was successfully employed by Baker, Harman and Pintz [1] in their result $d_n \ll p_n^{21/40}$. Employing a suitable sieve might enable one to avoid the critical case in our argument when $\sigma = 3/4$ and $x^{11/20} \le \tau \le x^{5/8}$, thereby enabling us to improve on the overall result.

Yu [32] employed a large double sieve to the problem when assuming the Lindelöf hypothesis. Although it appears that following exactly the method he employed does not improve the exponent when the Lindelöf assumption is dropped, the large double sieve could potentially aid the argument in another form. Following Yu's argument but using the bound $\zeta(1/2+it) \ll t^{\theta}$ gives a bound which approaches $x^{2+\epsilon}$ continuously as θ approaches 0. Using the best existing estimates of the order of $\zeta(1/2+it)$ (which are slightly smaller than 1/6) fails to produce an exponent better than $5/4$, and the argument does not seem to avoid the complications of the critical case in our argument.

References

[1] Baker, R. C., Harman, G., Pintz, J. The difference between consecutive primes, II. *Proc. London Math. Soc 83* (2001), 532–562.

[2] Bertrand, J. Mémoire sur le nombre de valeurs que peut prendre une fonction quand on y permute les lettres qu'elle renferme. *J. l'École Roy. Polytech. 17* (1845), 123–140.

[3] Chebyshev, P. Mémoire sur les nombres premiers. *Mém. Acad. Sci. St. Petersbourg 7* (1854), 17–33.

[4] Cramér, H. On the order of magnitude of the difference between consecutive prime numbers. *Acta Arith. 2* (1936), 23–46.

[5] de la Vallée Poussin, C.-J. Recherches analytiques sur la théorie des nombres premiers. *Ann. Soc. Scient. Bruxelles 20* (1896), 183–256.

[6] Erdős, P. The difference of consecutive primes. *Duke Math. J. 6* (1940), 438–441.

[7] Granville, A. Harald Cramér and the distribution of prime numbers. *Scandinavian Actuarial J. 1* (1995), 12–28.

[8] Hadamard, J. Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. *Bull. Soc. Math. France 24* (1896), 199–220.

[9] Heath-Brown, D. R. The twelfth power mment of the Riemann zeta-function. *Quart. J. Math. Oxford (2) 29* (1978), 443–462.

[10] Heath-Brown, D. R. The differences between consecutive primes. *J. London Math. Soc. (2) 18* (1978), 7–13.

[11] Heath-Brown, D. R. The differences between consecutive primes, II. *J. London Math. Soc. (2) 19* (1979), 207–220.

[12] Heath-Brown, D. R. The differences between consecutive primes III. *J. London Math. Soc. 20* (1979), 177–178.

[13] Heath-Brown, D. R. Zero density estimates for the Riemann zeta-function and Dirichlet L-functions. *J. London Math. Soc. (2) 19* (1979), 221–232.

[14] Heath-Brown, D. R. Prime numbers in short intervals and a generalized Vaughan identity. *Canad. J. Math. 34* (1982), 1365-1377.

[15] Heath-Brown, D. R. The number of primes in a short interval. *J. Reine Angew. Math. 389* (1988), 22–63.

[16] Heilbronn, H. Über den Primzahlsatz von Herrn Hoheisel. *Math. Z. 36* (1933), 394–423.

[17] Hoheisel, G. Primzahlprobleme in der Analysis. *Berliner Silzungsberichte* (1930), 580–588.

[18] Huxley, M. N. On the difference between consecutive primes. *Invent. Math. 15* (1972), 164-170.

[19] Ingham, A. E. Mean-value theorems n the theory of the Riemann zeta-function. *Proc. London Math. Soc. (2) 27* (1926), 273-300.

[20] Ingham, A. E. On the difference between consecutive primes. *Q. J. Math. 8* (1936), 255-266.

[21] Legendre, A. M. *Essai sur la théorie des nombres*. Paris Duprat, 1798.

[22] Montgomery, H. L. *Topics in multiplicative number theory*, vol. 227 of *Lecture notes in Math.* Springer, Berlin, 1971.

[23] Peck, A. S. *On the differences between consecutive primes*, DPhil thesis. University of Oxford, 1996.

[24] Pintz, J. Very large gaps between consecutive primes. *J. Number Theory 63* (1997), 286–301.

[25] Rankin, R. A. The difference between consecutive prime numbers, III. *J. London Math. Soc. 22* (1947), 226–230.

[26] Richert, H. Zur Abschätzung der Riemannschen Zetafunktion in der Nähe der Vertikalen $\sigma = 1$. *Math. Ann. 169* (1967), 97–101.

[27] Selberg, A. On the normal density of primes in small intervals and the difference between consecutive primes. *Arch. Math. Naturvid. 47* (1943), 87–105.

[28] Shanks, D. On maximal gaps between successive primes. *Math. Comput. 18* (1964), 646–651.

[29] Tchudakoff, N. G. On zeros of Dirichlet's L-functions. *Rec. Math. [Mat. Sbornik] N.S. 1 (43)* (1936), 591–602.

[30] Titchmarsh, E. C. *The theory of the Riemann zeta-function*, second ed. Oxford University Press, New York, 1986.

[31] Westzynthius, E. Über die Verteilung der Zahlen die zu den ersten Primzahlen teilerfremd sind. *Comm. Phys. Math. Helingsfors 5* (1931), 1–37.

[32] Yu, G. The differences between consecutive primes. *Bull. London Math. Soc. 28* (1996), 242–248.

Mathematical Institute, 24-29 St Giles', Oxford, OX1 3LB